

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM n cleic - nucleic search, using sw model

Run n: June 2, 2000, 08:56:29 ; Search time 5907.01 Seconds
(without alignments)
2750.176 Million cell updates/sec

Title: US-09-284-180-1
Percent score: 4008
Sequence: 1 gccgagccgcgcagtagc.....aaaaaaaaaaaaaaaaaaaaa.4008

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Search: 4857316 seqs, 2026611650 residues
Total number of hits satisfying chosen parameters: 9714632

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database:

EST:
1: em_est1:
2: em_est2:
3: em_est3:
4: em_est4:
5: em_est5:
6: em_est6:
7: em_est7:
8: em_est8:
9: em_est9:
10: em_est10:
11: em_est11:
12: em_est12:
13: em_est13:
14: em_est14:
15: em_est15:
16: em_est16:
17: em_est17:
18: em_est18:
19: em_est19:
20: gb_est1:
21: gb_est2:
22: gb_est3:
23: gb_est4:
24: gb_est5:
25: gb_est6:
26: gb_est7:
27: gb_est8:
28: gb_est9:
29: gb_est10:
30: gb_est11:
31: gb_est12:
32: gb_est13:
33: gb_est14:
34: gb_est15:
35: gb_est16:
36: gb_est17:
37: gb_est18:
38: gb_est19:
39: gb_est20:
40: gb_est21:
41: gb_est22:
42: gb_est23:
43: gb_est24:

44: gb_est25:
45: gb_est26:
46: gb_est27:
47: gb_est28:
48: gb_est29:
49: em_est20:
50: em_est21:
51: em_est22:
52: gb_est30:
53: gb_est31:
54: gb_est32:
55: em_est23:
56: em_est24:
57: em_est25:
58: em_est26:
59: gb_est33:
60: gb_est34:
61: gb_est35:
62: gb_est36:
63: gb_est37:
64: gb_est38:
65: em_est27:
66: em_est28:
67: em_est29:
68: em_est30:
69: gb_est39:
70: gb_est40:
71: gb_est41:
72: gb_est42:
73: gb_est43:
74: gb_est44:
75: em_est32:
76: em_est32:
77: em_est33:
78: em_est34:
79: gb_est45:
80: gb_est46:
81: gb_est47:
82: gb_gss1:
83: gb_gss2:
84: gb_gss3:
85: gb_gss4:
86: em_gss1:
87: em_gss2:
88: em_gss3:
89: em_gss4:
90: gb_gss5:
91: gb_gss6:
92: em_gss5:
93: em_gss6:
94: gb_gss7:
95: gb_gss8:
96: gb_gss9:
97: em_gss7:
98: em_gss8:
99: gb_gss11:
100: gb_gss10:
101: em_gss9:
102: em_gss10:
103: em_gss11:
104: em_gss12:
105: gb_gss12:
106: gb_gss13:
107: gb_gss14:
108: gb_gss15:
109: gb_gss16:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	Score	Query	Length	DB	ID	Description
1	437.2	10.9	466	69	AM123399	AM123399 UT-M-BH2.
2	405.2	10.1	434	69	AM123577	AM123577 UT-M-BH2.
3	386.8	9.7	503	34	AA459837	AA459837 zx05g12.r
4	293.6	7.3	517	22	RS4487	RS4487 yg74e12.r1
5	289.6	7.2	337	20	245329	245339 HSC2MG091.n
6	287.4	7.2	364	20	T09073	T09073 EST06966.in
7	284.4	7.1	376	81	AM436751	AM436751 77209.MAR
8	283.8	7.1	437	73	H24181	H24181 ym55f02.r1
9	281.6	7.0	347	22	H10623	H10623 ym08g07.r1
10	258	6.4	379	42	AM158284	AM158284 ud33h09.r
11	211.8	5.3	305	69	AM122335	AM122335 UT-M-BH2.
12	193.6	4.8	592	106	AO352219	AO352219 CTTB-ET1-
13	150.8	3.8	208	81	AM436702	AM436702 77141.MAR
14	138.8	3.5	257	72	AV329323	AV329323 AV329323
15	118.2	2.9	490	45	AI365650	AI365650 ap20h08.x
16	111.4	2.8	619	64	AL118624	AL118624 DFE2P761F
17	110.2	2.7	232	80	AM346693	AM346693 29273.MAR
18	110.2	2.7	232	80	AM346698	AM346698 29282.MAR
19	109	2.7	212	20	Z42773	Z42773 HSC0FE021.n
20	102.4	2.6	463	74	AM206442	AM206442 UT-H-B11-
21	100.8	2.5	558	44	AI113725	AI113725 HA335.Hu
22	99.2	2.5	395	44	AI289278	AI289278 gw22e06.x
23	98.6	2.5	359	20	T35785	T35785 EST91372.Hu
24	97.8	2.4	442	26	W39580	W39580 zc20c05.r1
25	94.8	2.4	350	20	Z43006	Z43006 HSC10E101.n
26	94.4	2.4	450	70	AM148766	AM148766 x102h03.x
27	87.8	2.2	427	47	AI533743	AI533743 th18h02.x
28	86.6	2.2	453	22	H08892	H08892 y188h02.r1
29	85	2.1	448	61	AI8588290	AI8588290 w135e08.x
30	84	2.0	591	34	AA532537	AA532537 n330e03.s
31	76.8	1.9	942	69	AU078986	AU078986 AU078986
32	75.6	1.9	442	48	AI555178	AI555178 UT-R-C2P-
33	75	1.9	638	79	AM245910	AM245910 2822888.5
34	74.6	1.9	395	26	W52192	W52192 zc46g02.s1
35	71.8	1.8	469	42	AI141990	AI141990 co017g10.x
36	69.4	1.7	313	38	AA761030	AA761030 in08h08.s
37	69.4	1.7	697	69	AM134401	AM134401 in18h09.Y
38	69	1.7	519	94	AO925155	AO925155 RCT1-23-3
39	69	1.7	737	109	AO645355	AO645355 RPT13-BEc
40	68	1.7	479	82	FR0003554	287337 F_rubirdlps
41	68	1.7	802	69	AM128036	AM128036 f106a12.Y
42	68	1.7	1104	82	CNS0041E	AL078534 drcosphi1
43	67.8	1.7	1203	82	CNS00747	AL063347 drcosphi1
44	67.4	1.7	600	83	FR0042485	AL129977 Fugu.rubr
45	67	1.7	449	47	AI545466	AI545466 f081902.x

RESULT	1
AM123399	
LOCUS	
DEFINITION	
ACCESSION	
VERSION	
KEYWORDS	
SOURCE	
ORGANISM	
REFERENCE	
AUTORS	
TITLE	
JOURNAL	
MEDIA	
COMMENT	

AM123399 466 bp mRNA EST 22-OCT-1999
 UT-M-BH2.1-apj-a-04-0-UI.s1 NIH.BMAP.M.S3.1 Mus musculus cDNA clone
 UT-M-BH2.1-apj-a-04-0-UI.3', mRNA sequence.
 AM123399
 AM123399.1 GI:6098929
 EST.
 house mouse.
 Mus musculus.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 466)
 Bonaldo,M.F., Lennon,G. and Soares,M.B.
 Normalization and subtraction: two approaches to facilitate gene
 discovery
 Genome Res. 6 (9), 791-806 (1996)
 97044477
 On Mar 10, 1998 this sequence version replaced gi:2948620.
 Contact: Chin, H

National Institute of Mental Health
 6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
 20892-9643, USA
 Tel: 301 443 1706
 Fax: 301 443 9890
 Email: MEST@mail.nih.gov
 Oligo-dT track not found. Not a site shown in beginning of sequence
 is likely internal to the message. cDNA Library Preparation: M.B.
 Scores Lab Clone distribution: NIH BMP cDNA clones will be made
 available by the means that is soon to be determined. When NIH
 determines the means for distribution of the BMP cDNA clones, this
 record will be updated accordingly when that means is determined.
 Seq primer: M13 Forward
 POLYA-No.

FEATURES	Location/Qualifiers
source	1. .466

/organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="U1-M-BH2.1-apj-a-04-0-UI"
 /clone_11b="NIH_BMAP_M.S3.1"
 /dev_stage="27-32 days"
 /lib_host="DH10B (Life Technologies)"
 /note="Vector: pT773D-Pac (Pharmacia) with a modified
 polylinker. Site_1: Not I; Site_2: Eco RI; The
 NIH_BMAP_M.S3.1 library is a subtracted library of a
 series, ultimately derived from a mixture of individually
 tagged normalized libraries from ten regions of the mouse
 brain (cerebellum, brain stems, olfactory bulbs,
 hypothalamus, cortex, amygdala, basal ganglia, pineal
 gland, striatum, hippocampus) after a series of
 subtractions to reduce the representation of cDNAs from
 which ESTs had already been generated. The following
 serially subtracted libraries were generated in this
 process: NIH_BMAP_M.S3.1, NIH_BMAP_M.S2, NIH_BMAP_M.S1.
 The subtracted library (NIH_BMAP_M.S3.1) was constructed
 as follows: PCR-amplified cDNA inserts from NIH_BMAP_M.S2
 clones from which 3' ESTs had been derived was used as a
 driver in a hybridization with the NIH_BMAP_M.S2 library
 in the form of single-stranded circles. The remaining
 single-stranded circles (subtracted library) was purified
 by hydroxyapatite column chromatography, converted to
 double-stranded circles and electroporated into DH10B
 bacteria (Life Technologies) to generate the
 NIH_BMAP_M.S3.1 library. This procedure has been
 previously described (Ronald, Lennon and Soares, Genome
 Research 6:791-806, 1996)
 TAG_lib=NIH_BMAP_M.S3.1
 TAG_tissue=pineal-glands
 TAG_SBO=CGAC"

BASE COUNT	95 a	139 c	120 g	112 t
ORIGIN				

Query match	Score	DB	Length
10.98;	437.2;	69;	466;

Matches 448; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

189 **cgggcgcgtcccccgtcagtgcccagaacctcgctgcccatctccgaggtgactccta** 248

1 CGGCCGCTCCCCGCTCGGTGCCCAAGAACCTCGCTGCCCATCTCCGAGGCTGACTCCTA 60

249 tctcaccgcggttcgcagcgcgtctcatacgtacaaattactctgcctccctgttgatccctgc 308

01 TCTACCCGGTTGGGCCCTTCAIACGTAACAATATACITCGTCTCTCCITGGTGGATCTGC 120

[illegible]

T

[illegible]

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

```

OY 429 gaagaaagcgaagaagagcgaatgcacatttaccagattctcgccatgtcaa 488
DB 241 GAAGAAAGCAGAAAGAGAGAGATGTACAAATTTATTCAGATTCTCGCATGGCCAA 300
OY 489 tgccttaccctctcctcagctcgcagccttcgtttatccgaagtcggggtattga 548
DB 301 TGCTCTACCTCTCTCAGCTGTGGACCTTGCTTTTATCCGAAGTGGGGGTATTGA 360
OY 549 tgttccagttccagcagcgtttgaaagacttgagagcgccgggggaaatctcttga 608
DB 361 TGTGTCCAGTTTCCAGCAGGTTGAAAGACTTGAGAGTGGCGGGGGAATGCTTTGA 420
OY 609 gccagctcaacgctcagcagctgtatgtcgtgggggctcctctac 654
DB 421 GCCAGCTCAACGCTCAGCAGCTGTATGTGCTGGGGGCGTCTCTAC 466

RESU T 2
AM12 577 AM123577 434 bp mRNA EST 22-OCT-1999
LOCUS UI-M-BH2.1-ap1-a-04-0-UI.S1 NIH_BMAP_M.S3.1 Mus musculus cDNA clone
DEFINITION UI-M-BH2.1-ap1-a-04-0-UI 3', mRNA sequence.
ACCESSION AM123577
VERSION AM123577.1 GI:6099107
KEYWORDS EST.
SOURCE house mouse.
ORIGIN Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Scuriognathi; Muridae; Murinae; Mus.
1 (bases 1 to 434)
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
97044477
On Mar 10, 1998 this sequence version replaced gi:2948806.
Contact: Chin, H
National Institute of Mental Health
6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
20892-9643, USA
Tel: 301 443 1706
Fax: 301 443 9890
Email: MEST@mail.nih.gov
Oligo-dt track not found, Not I site shown in beginning of sequence
is likely internal to the message. cDNA Library Preparation: M.B.
Soares Lab clone distribution: NIH BMAP cDNA clones will be made
available by the means that is soon to be determined. When NIH
determines the means for distribution of the BMAP cDNA clones, this
record will be updated accordingly when that means is determined.
Seq primer: M13 Forward
POLY-A-NO.
Location/Qualifiers
1. 434
/organism="Mus musculus"
/strain="CS7BL/6J"
/db_xref="taxon:10090"
/clone="UI-M-BH2.1-ap1-a-04-0-UI"
/dev_stage="27-32 days"
/lab_host="DH10B (Life Technologies)"
/note="Vector: pUT3D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; The
NIH_BMAP_M.S3.1 library is a subtracted library of a
series, ultimately derived from a mixture of individually
tagged normalized libraries from ten regions of the mouse
brain (cerebellum, brain stems, olfactory bulbs,
hypothalamus, cortex, amygdala, basal ganglia, pineal
gland, striatum, hippocampus) after a series of
subtractions to reduce the representation of cDNAs from
which ESTs had already been generated. The following
serially subtracted libraries were generated in this
process: NIH_BMAP_M.S3.1, NIH_BMAP_M.S2, NIH_BMAP_M.S1.

```

```

BASE COUNT 90 a 130 c 109 g 105 t
ORIGIN
Query Match 10.1%; Score 405.2; DB 69; Length 434;
Best Local Similarity 95.9%; Pred. No. 1.4e-81;
Matches 416; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

OY 189 cggccgcgtcccccgcgcagtgcccaagaacctgcgcacatcccgaggtgactccta 248
DB 1 CGGCGCGCTCCCGCGCTCGTGCCCAAGACTCCGCTGCCATCTCCAGAGCTGACTCCTA 60
OY 249 tctaccgggttgcagcgtctcagctacagtaacttctccttctgtgactctgc 308
DB 61 TCTCACCCTGTTGCGGCCCTCATAGCTACATTAATCTCTCTCTGTTGATCTCTGC 120
OY 309 ctccacacacttaacgtcgtgtcagggagataagatcttgccttaaccctcccttc 368
DB 121 TTCACACACTTAACTGCGGTGACGACGACGACATCTCCCTTAACTCCCTCTTC 180
OY 369 tggggaaagaccccgaaagatcgaactgatatgtactctctagactacagaagaactgag 428
DB 181 TGGGAAAAGCTCTGAAAGATCGACTGATGTGTGCCAGACTACAGACGAAGCTGTAG 240
OY 429 gaagaaagcgaagaagagcgaatgtcacattttaccagattctcgccatgtcaa 488
DB 241 GAAGAAAGCAGAAAGAGAGAGATGTACAAATTTATTCAGATTCTCGCATGGCCAA 300
OY 489 tgccttaccctctcctcagctcgcagccttcgtttatccgaagtcggggtattga 548
DB 301 TGCTCTACCTCTCTCAGCTGTGGACCTTGCTTTTATCCGAAGTGGGGGTATTGA 360
OY 549 tgttccagttccagcagcgtttgaaagacttgagagcgccgggggaaatctcttga 608
DB 361 TGTGTCCAGTTTCCAGCAGGTTGAAAGACTTGAGAGTGGCGGGGGAATGCTTTGA 420
OY 609 gccagctcaacgct 622
DB 421 GCCAGCTCAACGCT 434

RESULT 3
LOCUS AA459837 503 bp mRNA EST 09-JUN-1997
DEFINITION x550q12.1 Soares testis NHT Homo sapiens cDNA clone IMAGE:795718
5' similar to TR:G554328 G854328 SEAHPRORIN C, mRNA sequence.
ACCESSION AA459837
VERSION AA459837.1 GI:2184744
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiinae; Homo.
REFERENCE 1 (bases 1 to 503)
Haller, L., Allen, M., Bowles, L., Dubuque, J., Gelsel, G., Jost, S.,
Kucaba, T., Lacy, M., Lennon, G., Marra, M., Martin, J.,
Moore, B., Schellenberg, K., Stepien, M., Tan, F., Thelning, B.,
White, Y., Wylie, T., Waterston, R. and Wilson, R.

```

The subtracted library (NIH_BMAP_M.S3.1) was constructed as follows: PCR amplified cDNA inserts from NIH_BMAP_M.S2 clones from which 3' ESTs had been derived were used as a driver in a hybridization with the NIH_BMAP_M.S2 library in the form of single-stranded circles. The remaining single-stranded circles (subtracted library) was purified by hydroxyapatite column chromatography, converted to double-stranded circles and electroporated into DH10B bacteria (Life Technologies) to generate the NIH_BMAP_M.S3.1 library. This procedure has been previously described (Bonaldi, Lennon and Soares, Genome Research 6:791-806, 1996)

TAG: NIH_BMAP_M.S3.1
TAG-TISSUE: pineal glands
TAG-SEQ-CCATC

WASHU-MERCK EST Project 1997
 Unpublished (1997)
 On Sep 12, 1996 this sequence version replaced gi:1392936.
 Contact: Wilson RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: estewatson.wustl.edu
 This clone is available royalty-free through LNLN; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Seq primer: -28ml3 rev2 ET from AmerSham
 High quality sequence stop: 459.

FEATURES

source

1. 503
 /organism="Homo sapiens"
 /db_xref="GDB:6039163"
 /db_xref="taxon:9606"
 /clone="IMAGE:795718"
 /clone_lib="Soares_testis_NH"
 /sex="male"
 /lab_host="DH10B"
 /note="Vector: pT73D-Pac (Pharmacia) with a modified
 polylinker. Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
 was prepared from mRNA obtained from Clontech
 Laboratories, Inc., and primed with a Not I - oligo(dT)
 primer [5].
 TGTACCAATCTGAACTGAGGAGGCGGCCCAATTTTCTTTTCTTTT 3']
 Double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Not I and cloned into the Not I
 and Eco RI sites of the modified pT73 vector. Library
 went through one round of normalization to Cot5, and was
 constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 105 a 136 c 149 g 113 t
 ORIGIN

Query Match 9.7%; Score 386.8; DB 34; Length 503;
 Best Local Similarity 88.2%; Pred. No. 2.2e-77;
 Matches 443; Conservative 0; Mismatches 57; Indels 2; Gaps 2;

1387 gactatcagagtcgtggtccacaggtgtaccagcctctcagggaaagataagctg 1446
 1 GCCTATCTCAGAGTCGTGGCCACA -GGTACACAGCCTCTCAGGAAGATGATGCTG 59
 1447 ccttaacctggagaca-gaagatggacacctccacgggtctgtgcacatggagctacgt 1505
 60 CTCTACTCTGGGACATAGAGATGACACCTCCACCGAGAGTGGGATCGAGCTCAGCT 119
 1506 cagtgtcttggaagatctgtctgtccagagacacacagcggttagagagcatgaatt 1565
 120 CACCGTTCTTGAAGATCTGGCTTATTTCCAGAGCCACACCGATTGAGAACATGAAATT 179
 1566 gtacacagatgtgctcctgtgtggtcccaactagaggtgacacaaagtacacacagca 1625
 180 GTACACAGATGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 229
 1626 ctgtgacgctctccagagatctgtcgaagtgatccgtggtccacagacccgtgtgcgctg 1685
 240 CTGTGCGCGCTCTCAGAGCTGCTCAGAGTGCATCTGCGCCAGACCCCACTGTGCTG 299
 1686 gaagctcggagctgagctgtgtgtgcccagcgcgagagacacggagatgttcaaga 1745
 300 GACCTTCGGCTGGATGAGTGTGGCCCATGCGGGGAGACCGAGAGGTGGTCCAGAGA 359
 1746 taagagtcagcagatgctctctctgtgtccaaagaacccctggaacatcccgtagt 1805
 360 CAAAGATGACAGATGCT 419
 1806 gtttgaagtcgggtgctactgtgtggcacaagtgtcctgtccatgttccccagttctgc 1865
 420 GTTTGAAGTCCCGTGGCTACAGCTCGGCATGTGTTCTTCCCATGTCTCCAGACTCAGC 479

QY 1866 ctggacatcctgtgtgtgac 1887
 Db 480 ATGGCATCCTGTGTGGCAC 501

RESULT 4
 LOCUS R54387
 DEFINITION y978f12.r1 Soares infant brain IN1B Homo sapiens cDNA clone
 IMAGE:39491 5' similar to SP:A49069 A49069 COLLA5P1 - ; mRNA
 sequence.

ACCESSION R54387
 VERSION R54387.1 GI:816289
 EST.
 KEYWORDS EST.

SOURCE

ORGANISM

human.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE

AUTHORS

Hallier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,
 Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marita, M.,
 Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F.,
 Trevasakis, E., Waterston, R., Williamson, A., Woldmann, P. and
 Wilson, R.

TITLE

JOURNAL

COMMENT

The WashU-Merck EST Project
 Unpublished (1995)
 Contact: Wilson RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: estewatson.wustl.edu
 Insert Size: 2134
 High quality sequence stops: 99 Source: IMAGE Consortium, LNLN. This
 clone is available royalty-free through LNLN; contact the IMAGE
 Consortium (info@image.llnl.gov) for further information.
 Insert Length: 2134 Std Error: 0.00
 Seq primer: M13Rpi
 High quality sequence stop: 99.

FEATURES

source

1. 517
 /organism="Homo sapiens"
 /db_xref="GDB:412032"
 /db_xref="taxon:9606"
 /clone="IMAGE:39491"
 /clone_lib="Soares infant brain IN1B"
 /sex="female"
 /dev_stage="73 days post natal"
 /lab_host="DH10B (ampicillin resistant)"
 /note="Organ: whole brain; Vector: Lambda B4; Site 1: Not
 I; Site 2: Hind III; 1st strand cDNA was primed with a Not
 I - oligo(dT) primer [5].
 AACGTGAAGAAATTCGGCGCGCGCGGGAATTTTCTTTTCTTTT 3']
 double-stranded cDNA was ligated to Hind III adaptors
 (Pharmacia), digested with Not I and directionally cloned
 into the Not I and Hind III sites of the Lambda B4 vector.
 Library went through one round of normalization. Library
 constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT

ORIGIN

Query Match 7.3%; Score 293.6; DB 22; Length 517;
 Best Local Similarity 85.6%; Pred. No. 2.9e-56;
 Matches 370; Conservative 0; Mismatches 55; Indels 7; Gaps 4;

QY 1127 gaacccaagacatcggagagtcgaatggtcccttagagagctaaacatgactga 1186
 Db 1 GACCAACAGACATTCGACAGAGTGTCTTCAAGAGACTTAACATGACTGCA 60
 QY 1187 acagggagactgctgcatgagacaagaggtgcccagccagacctgagagtgatcg 1246
 Db 61 ACAGAGAGTGGCTGTGCGTGAATGATGTGGCCAGCCAGAGACTGAGATGATCA 120

QY 1247 ccaacaacatgaagccagcagcttgatctcctcctccagacccgctgctca 1306
|||||
DB 121 ccaacacatgaagctccgacattgctcctcctcctcctcctcctcctcctc 180
QY 1307 ccttatacagaacacccctcctcctcctcctcctcctcctcctcctcctc 1366
|||||
DB 181 ccttatacagaacacccctcctcctcctcctcctcctcctcctcctcctc 240
QY 1367 tctgtgctacatacagatcagctcctcctcctcctcctcctcctcctcctc 1426
|||||
DB 241 tctgtgctacatacagatcagctcctcctcctcctcctcctcctcctcctc 300
QY 1427 cagggaagaataatgagctgctcctcctcctcctcctcctcctcctcctc 1484
|||||
DB 301 cagggaagaataatgagctgctcctcctcctcctcctcctcctcctcctc 360
QY 1485 tctgtgctacatacagatcagctcctcctcctcctcctcctcctcctcctc 1539
|||||
DB 361 agtgcgcatcgcagctcagctcagctcagctcagctcagctcagctcagctc 420
QY 1540 ccacacccgctt 1551
|||||
DB 421 tcacacnccagtt 432
RESU T 5
245329 337 bp mRNA EST 14-NOV-1994
LOCUS HSCMG091 normalized infant brain cDNA Homo sapiens cDNA clone
DEFINITION c-2mg09, mRNA sequence.
ACCESSION 245329
VERSION 245329.1 GI:574541
KEYWORDS EST.
SOURCE human.
ORANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
1 (bases 1 to 337)
Auffray, C., Behar, G., Bols, F., Bouchier, C., da Silva, C.,
Devignes, M.D., Dupret, S., Houliatte, R., Jumeau, M.N., Lam, B.,
Lorenzo, F., Mitchell, H., Marigault-Samson, R., Pietu, G., Pouliot, Y.,
Sebastien, K., Kachichis, C., and Tessier, A.
IMAGE: molecular integration of the analysis of the human genome
and its expression
C. R. Acad. Sci., III, Sci. Vie 318 (2), 263-272 (1995)
95277534
CONTACT: Genethon
Genethon
Genethon Centre de recherche sur le Genome Humain
1, rue de l'Internationale, BP60 91002 EVRY Cedex, FRANCE
Tel: 33169472800
Fax: 33160778698
Email: genexpress@genethon.fr
Single read
Genexpress_library_id: C; Genexpress_sequence_id: y1c-2mg09
Seq primer: (-21)M13-universal.
Location/Qualifiers
1..337
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="c-2mg09"
/clone_1lb="normalized infant brain cDNA"
/sex="Female"
/tissue_type="total brain"
/dev_stage="3 months old"
/note="Organ: brain; Vector: lafmid BA. Site 1: HindIII;
Site 2: NotI; sex=Female; dev_stage=3 months old;
Isolate=muscular atrophy patient; tissue_type=total
brain; total mRNA was oligo-(dT) primed and directionally
cloned 5' -> 3' into the HindIII -> NotI sites of the
lafmid BA vector. Clone library from B.Soures, Psychiatry
Dept. Columbia University, USA. Normalization_method:

BASE COUNT 81 a 106 c 82 g 67 t 1 others
ORIGIN
Query Match 7.2% Score 289.6; DB 20; Length 337;
Best Local Similarity 91.1%; Pred. No. 2e-55;
Matches 307; Conservative 0; Mismatches 30; Indels 0; Gaps 0;
QY 1126 ccacccaagaatccagcagctgctgaaagctcctttagagagctaaacatgctgc 1185
|||||
DB 1 ccacccaagaatccagcagctgctgaaagctcctttagagagctaaacatgctgc 60
QY 1186 aacagggagctcctgctgctgctgctgctgctgctgctgctgctgctgctg 1245
|||||
DB 61 aacagggagctcctgctgctgctgctgctgctgctgctgctgctgctgctg 120
QY 1246 gccacaacatgaagctcagcagctgctgctgctgctgctgctgctgctgctg 1305
|||||
DB 121 gccacaacatgaagctcagcagctgctgctgctgctgctgctgctgctgctg 180
QY 1306 accctatcagaagacacccctcctcctcctcctcctcctcctcctcctcctc 1365
|||||
DB 181 accctatcagaagacacccctcctcctcctcctcctcctcctcctcctcctc 240
QY 1366 ctgctgctacatacagatcagctcctcctcctcctcctcctcctcctcctc 1425
|||||
DB 241 ctgctgctacatacagatcagctcctcctcctcctcctcctcctcctcctc 300
QY 1426 tcagggaagaataatgagctgctcctcctcctcctcctcctcctcctcctc 1462
|||||
DB 301 tcagggaagaataatgagctgctcctcctcctcctcctcctcctcctcctc 337
RESULT 6
LOCUS T09073 364 bp mRNA EST 03-AUG-1993
DEFINITION EST06966 Infant Brain, Bento Soares Homo sapiens cDNA clone HIB071
5' end, mRNA sequence.
ACCESSION T09073
VERSION T09073.1 GI:390101
KEYWORDS EST.
SOURCE human.
ORANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
1 (bases 1 to 364)
Adams, M.D., Soares, M.B., Kerlavage, A.R., Fields, C. and Venter, J.C.
Rapid cDNA sequencing (expressed sequence tags) from a
directionally cloned human infant brain cDNA library
Nature Genet. 4, 373-380 (1993)
94004965
CONTACT: Adams, MD
The Institute for Genomic Research
932 Clopper Road, Gaithersburg, MD 20878
Tel: 3018699096
Fax: 3018699423
Email: mdadams@tigr.org
Seq primer: M13 Reverse.
Location/Qualifiers
1..364
/organism="Homo sapiens"
/db_xref="ATCC (Inhost):85555"
/db_xref="taxon:9606"
/clone="HIB071"
/clone_1lb="Infant Brain, Bento Soares"
FEATURES
Source
BASE COUNT 70 a 100 c 106 g 85 t 3 others
ORIGIN
Query Match 7.2% Score 287.4; DB 20; Length 364;
Best Local Similarity 86.5%; Pred. No. 6.5e-55;
Matches 315; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

QY 1561 aaattgtaccagcattggtctctctgtgtggtccctactagtgtgacacaaagttaacacc 1620
 |||
 Db 1 AAATTGTACACAGACTGCTCTCTGTTGGCTCCGTTACTGTAGTGACACAGTGAATACA 60
 QY 1621 agcaactgtgacgtctcccaagactgtctgagatgtatccctggccagagaccctgtgac 1680
 |||
 Db 61 ACNACTGTGGCCCTCTCCAGAGCTGCTCAGAGTGATCCCTGGCCAGACCCAGTCTGT 120
 QY 1681 gccctgagctcccgctgtgagctgtgtgtgcccacgcgcgagacacgcggagtgtt 1740
 |||
 Db 121 GCCGTGAGCTCCCGCTGTAGTAGAGTGTGGCCATGCCGGGAGACACGAGGGTGTGTC 180
 QY 1741 caagatatagagtcagcagcagatgtctctctgtgtgtccaaagaacatgtgagacatccc 1800
 |||
 Db 181 CAACACATAGATGATGACGAGATGTCTCTCTTGTGTCTTAAAGCCCTGGAGAACGTGCA 240
 QY 1801 gtatgtttgaagttccggt 1860
 |||
 Db 241 GTAGTGTGTGAAGTTCCTCGTGTGCTACAGNTGCGATGTGTGTCTTNCATGTCTTCCAAAGC 300
 QY 1861 tctgcttgagcactcctgt 1920
 |||
 Db 301 TCAGCATGAGGATCTCTGT 360
 QY 1921 aggg 1924
 |||
 Db 361 CGGG 364

RESULT 7

AM436751 376 bp mRNA EST 14-FEB-2000
 LOCUS 77209 MARC 2Pig Sus scrofa cDNA 5', mRNA sequence.
 DEFINITION AM436751
 ACCESSION AM436751
 VERSION AM436751.1 GI:6972057
 KEYWORDS EST.
 SOURCE pig.
 ORGANISM Sus scrofa
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Cetartiodactyla; Suidae; Sus.
 1 (bases 1 to 376)
 Fahnenkrug, S.C., Freking, B.A., Rohrer, G.A., Smith, T.P.L., Casas, E.,
 Stone, R.T., Heaton, M.P., Grosche, W.M., Bennett, G.A. and Keele, J.W.
 Design and use of two pooled tissue normalized cDNA libraries for
 EST discovery in swine
 Unpublished (2000)
 On Jul 9, 1999 this sequence version replaced gi:5435515.
 CONTACT Smith TPL
 USDA, ARS, US Meat Animal Research Center
 PO Box 166, Clay Center, NE 68933-0166, USA
 Tel: 402 762 4366
 Fax: 402 762 4366
 Email: smitht@mail.marc.usda.gov
 Single pass sequencing. Bases called and trimmed with phred
 v0.980904.e. Vector identified by cross_match with the -minscore 20
 and -mismatch 12 options.
 PCR primers
 FORWARD: AGGAACAGCTATGACCAT
 BACKWARD: GTTTTCCAGTCCAGCAGC
 Plate: 32 Row: A Column: 20
 Seq primer: ATTTAGGTGACACTATAG.
 Location/Qualifiers
 1..376
 /organism="Sus scrofa"
 /db_xref="taxon:9823"
 /clone_id="MARC 2Pig"
 /tissue_type="pooled"
 /lab_host="DH10B"
 /note="Vector: pCMV SPORT6, site_1: XbaI; site_2: XhoI;
 library made from pooled tissue from testis, ovary,
 endometrium, hypothalamus, pituitary, and placenta."

BASE COUNT

60 a 129 c 116 g 71 t

ORIGIN

Query Match 7.1%; Score 284.4; DB 81; Length 376;
 Best Local Similarity 86.9%; Pred. No. 3,1e-54;
 Matches 325; Conservative 0; Mismatches 46; Indels 3; Gaps 1;

QY 2032 agccaagcgggagacccttaaacccggcccaacacgctgtgtgtgtgtgtgtgtgtgtgt 2091
 |||
 Db 2 AGCCAGAGGGGGGCCCCCAACCCGGGCTCACAC--AGTGGGGCCGAGCTGAGAGGCTTC 58
 QY 2092 ctctcgtgtgtctctgtgagatccctcactcctcctcctgtgtgtgtgtgtgtgtgtgtgt 2151
 |||
 Db 59 TTCTGGGGGTTCTTGAGAGATCCCTGACTCTCTCTGTATGTGTGGCTCAGACGG 118
 QY 2152 cggcagcagcaggagcctctctagctagagacagcagcagcagcagcagcagcagcagc 2211
 |||
 Db 119 CAGGACAGAGGAGGCTTGTGCTAGAGACAGTGGGCTTGAGCTGGGGGTCACAGCA 178
 QY 2212 tctgtgaccccaagctatagtcagagaccctcctcctcctcctcctcctcctcctcct 2271
 |||
 Db 179 TCGGGACACAGCTACAGCCAGGACCTCCTCTCTCCCGAAGTAGCGGCTG 238
 QY 2272 cccctgacctgt 2331
 |||
 Db 239 CCCTGGCCCTGGCCAGAGAGGGGAGTGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 298
 QY 2332 gattcttgcacaaagccagccacatccggtcactgtgtgtgtgtgtgtgtgtgtgtgtgt 2391
 |||
 Db 299 GATCCTTGCCCGAGCCAGCCACATTCGCTGACGGGGGCTCCTGTAGCACATGTGAT 358
 QY 2392 gagacccatcata 2405
 |||
 Db 359 GAACGTCATCTA 372

RESULT 8

H24181 437 bp mRNA EST 06-JUL-1995
 LOCUS H24181
 DEFINITION ym55f02.r1 Soares infant brain JNIB Homo sapiens cDNA clone
 IMAGE:52280 5' similar to SP:A49069 A49069 COLLAPSin - ; mRNA
 sequence.
 ACCESSION H24181 GI:892876
 VERSION H24181
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 437)
 Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,
 Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,
 Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F.,
 Trevasakis, E., Waterston, R., Williamson, A., Woldmann, P. and
 Wilson, R.
 The WashU-Werck EST Project
 Unpublished (1995)
 CONTACT: Wilson RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@wustl.wustl.edu
 Insert Size: 1989
 High quality sequence stops: 310 Source: IMAGE Consortium, LNL
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.lnl.gov) for further information.
 Insert Length: 1969 Std Error: 0.00
 Seq primer: MJ3P1
 High quality sequence stop: 310.
 Location/Qualifiers
 1..437
 /organism="Homo sapiens"

FEATURES

source

/db_xref="GDB:425216"
 /db_xref="taxon:9606"
 /clone_image="52280"
 /clone_lib="Soares Infant brain INIB"
 /sex="female"
 /dev_stage="73 days post natal"
 /lab_host="DH10B (ampicillin resistant)"
 /note="Organ: whole brain; Vector: Lambda B; Site: 1: Not I; Site: 2: Hind III; 1st strand cDNA was primed with a Not I - oligo(dT) primer (5' AACTGGAAGATTCGGCGCCGACAGATTTTATTTTATTTT 3'); double-stranded cDNA was ligated to Hind III adaptors (Pharmacia), digested with Not I and directionally cloned into the Not I and Hind III sites of the Lambda B vector. Library went through one round of normalization. Library constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 105 a 122 c 112 g 95 t 3 others
 ORIG N

Query Match 7.1%; Score 283.8; DB 23; Length 437;
 Best Local Similarity 88.8%; Pred. No. 4.5e-54;
 Matches 318; Conservative 0; Mismatches 38; Indels 2; Gaps 1;

Db 1127 gaccccaagacatccggcagtcgtgaatggtcccttagagagctaaacatgactgca 1186
 |||||
 Db 1 GACCCAGAGACTTGTGGAGAGTGTGAATGGTCCCTTCAGAGAACTAAACATGACTGCA 60
 QY 1187 acaaggagactgctgtcatgacaagagtgccccagaccagactgagagtgatcag 1246
 |||||
 Db 61 ACAGAGAGACTGCTGTCTGTGGACATGATGTGCCACGCCAGACTGAGAGTGCATCA 120
 QY 1247 ccaacaacatgaagctccagcagttgtgactcctcctccagacgcggtgctca 1306
 |||||
 Db 121 CCACACATGATGAAGCTCCGGACATTTGGCTCATCTCTCCCTGCTGACCCGACTCA 180
 QY 1307 ccttatcagagacaccccttcacgagcagccggtgtcccgctgacggccgcccc 1366
 |||||
 Db 181 CCTTATCCGGGACACACCACTCATGAGAGCCAGTGTTCACCTGTGAGCCACCCCC 240
 QY 1367 tgcgtgactacacagatacagcctatctcaagatcgtgtgccccagaggtgacagcctc 1426
 |||||
 Db 241 TGCTGTCTACTACATACAGCTTATCTCAGAGTGTGCGCCACAGGGTGCACACCTTT 300
 QY 1427 cagggaagaataatgactgtctc--acctggagacagagatgacacctccacgg 1482
 |||||
 Db 301 TAGGGAAGAGATGATGATGTGCTTACCTTACCTGGGGACAGAGATGAGACACTTCCACCG 358

RESU T 9
 H106 3
 LOCUS 347 bp mRNA EST 23-JUN-1995
 DEFINITION ym8907.r1 Soares Infant brain INIB Homo sapiens cDNA clone
 IMAGE:47298 5' similar to SP:A49069 A49669 COLLAPSE -; mRNA
 sequence.

ACCE SION H10623
 VERS ON H10623.1 GI:875445
 KEYW RDS
 SOUR E
 OR ANISM Homo sapiens

REF ENCE
 AU HORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Hominoidea; Homo.

TI LE
 JO: RNAL
 COMM NT
 The WashU-Merck EST Project
 Unpublished (1995)
 On May 10, 1995 this sequence version replaced gi:805665.
 Contact: Wilson RK
 Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@wustl.wustl.edu
 Insert Size: 1985
 High quality sequence stops: 198
 Source: IMAGE Consortium, LNL
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.lnl.gov) for further information.
 Insert length: 1985 Std Error: 0.00
 Seq primer: M13RP1
 High quality sequence stop: 198.

FEATURES

source

location/qualifiers

1..347

/organism="Homo sapiens"

/db_xref="GDB:419839"

/db_xref="taxon:9606"

/clone_image="47298"

/clone_lib="Soares Infant brain INIB"

/sex="female"

/dev_stage="73 days post natal"

/lab_host="DH10B (ampicillin resistant)"

/note="Organ: whole brain; Vector: Lambda B; Site: 1: Not I; Site: 2: Hind III; 1st strand cDNA was primed with a Not I - oligo(dT) primer (5' AACTGGAAGATTCGGCGCCGACAGATTTTATTTTATTTT 3'); double-stranded cDNA was ligated to Hind III adaptors (Pharmacia), digested with Not I and directionally cloned into the Not I and Hind III sites of the Lambda B vector. Library went through one round of normalization. Library constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT

82 a 105 c 87 g 67 t 6 others

ORIGIN

Query Match 7.0%; Score 281.6; DB 22; Length 347;
 Best Local Similarity 90.1%; Pred. No. 1.3e-53;
 Matches 299; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 1127 gaccccaagacatccggcagtcgtgaatggtcccttagagagctaaacatgactgca 1186
 |||||
 Db 1 GACCCAGAGACTTGTGGAGAGTGTGAATGGTCCCTTCAGAGAACTAAACATGACTGCA 60
 QY 1187 acaaggagactgctgtcatgacaagagtgccccagaccagactgagagtgatcag 1246
 |||||
 Db 61 ACAGAGAGACTGCTGTCTGTGGACATGATGTGCCACGCCAGACTGAGAGTGCATCA 120
 QY 1247 ccaacaacatgaagctccagcagttgtgactcctcctccagacgcggtgctca 1306
 |||||
 Db 121 CCACACATGATGAAGCTCCGGACATTTGGCTCATCTCTCCCTGCTGACCCGACTCA 180
 QY 1307 ccttatcagagacaccccttcacgagcagccggtgtcccgctgacggccgcccc 1366
 |||||
 Db 181 CCTTATCCGGGACACACCACTCATGAGAGCCAGTGTTCACCTGTGAGCCACCCCC 240
 QY 1367 tgcgtgactacacagatacagcctatctcaagatcgtgtgccccagaggtgacagcctc 1426
 |||||
 Db 241 TGCTGTCTACTACATACAGCTTATCTCAGAGTGTGCGCCACAGGGTGCACACCTTT 300
 QY 1427 cagggaagaataatgactgtctc--acctggagacagagatgacacctccacgg 1458
 |||||
 Db 301 CAGGGAAGAGATGATGATGTGCTTACCTTACCTGGGG 332

RESULT 10

A1158284

LOCUS 379 bp mRNA EST 30-SEP-1998

DEFINITION u423h09.r1 Soares thymus_2NBM7 Mus musculus cDNA clone

IMAGE:1446785 5' mRNA sequence.

ACCESSION A1158284

VERSION A1158284.1 GI:3686753

KEYWORDS EST.

SOURCE house mouse.

	ORIGINISM	Mus musculus
		Eumetazoa; Chordata; Craniata; Vertebrata; Mammalia;
		Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE		1 (bases 1 to 379)
AUT HORS		Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Giesel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Thasing,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterson,R.
TITLE		The WashU-HMNI Mouse EST Project
JCI RNAL		Unpublished (1996)
COMMENT		On Jan 19, 1998 this sequence version replaced gi:2282400. Contact: Marra M/Mouse EST Project WashU-HMNI Mouse EST Project Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: mouseest@watson.wustl.edu This clone is available royalty-free through LNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information. MG:920101 Seq primer: -28m13 rev2 ET from Amersham High quality sequence stop: 373.
FEATL RES		Location/Qualifiers
SOURCE		1..379 /organism="Mus musculus" /strain="C57BL/6J" /db_xref="taxon:10090" /clone_IMAGE:1446785" /clone_1lb="Soares_1thymsu_2nbwt" /sex="male" /tissue_type="Thymus" /dev_stage="4 weeks" /lab_host="DH10B" /note="vector: pT73D-Pac (Pharmacia) with a modified polylinker. Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5'] TGTACCACTGTCAGTGAGCGGGCCGGTTTCTTTTTTTTTTTTTT [3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. RNA provided by Dr. Bertrand Jordan. Library went through two rounds of normalization, and was constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT		64 a 111 c 109 g 95 t
ORIGIN		
Que ry Match		6.4%; Score 238; DB 42; Length 379;
Bee c Local Similarity		90.9%; Pred. No.3e-48; 25; Indels 8; Gaps 5;
Mat ches 331; Conservative		0; Mismatches
Oy 2768	tctctcccaactcgaattcccccttg-ctctgaaagagtgtgtctatcacctgtgtgc	2826
Dd 11	ttcttcgccacctgatttccccttttgtcccttgaaagagtgtgtctatcacctgtgtgc	70
Oy 2827	ctagaaggcctgtcatgtgtcatgtacgtacgaagggcggtgctcgtgtcttttgggga	2886
Dd 71	ctcgaagacctgtccgctgtgtgatgacagggccggtgcctgtgtgttggggg	130
Oy 2887	gtcggagagaaaaggtgtgaaatgggggacaacttaacccttcggttagcagatgggaaacc	2946
Dd 131	gtcggagagaaaggtgtgaaatgggggacaacttaacccttcggttagcagatgggaaacc	190
Oy 2947	acatgcccgtcccatcacccacagcggtctttaactttagccaagtcccaaatg	3006
Dd 191	a-----cgcgaccccatcacccacagcggtctttaactttagccaagtcccaaatg	246
Oy 3007	acctctcgggtgggaaggagcagagacatgtgccccgcctctctcttcttccce	3066
Dd 247	acctctcgt-ggtgggaaggagcagagacacggg--cccgctccttctctcttcttctc	304

Oy	3067	ttctgcgtccacacatcgcgttcaccagctgcttcccttgctgaatgtagctg	3126
Dd	305	CTGTGGCGTCC-ACCACATGCCGTGCACACGCTGCTTTTCCTCCTGACTGATGAGAGCTG	363
Oy	3127 agtc	3130	
Dd	364 AGCC	367	
RESULT	11		
LOCUS	AM122335/c		
DEFINITION	305 bp mRNA	EST	22-OCT-1998
ACCESSION	UI-M-BH2.2-aos-g-02-0-UI.s1 NIH_BMAP_M.S3.2	Mus musculus cDNA clone	
VERSION	UI-M-BH2.2-aos-g-02-0-UI 3'	mRNA sequence.	
KEYWORDS	AM122335		
SOURCE	AM122335.1 GI:6097820		
ORGANISM	EST.		
	house mouse.		
	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;		
AUTHORS	Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
TITLE	1 (bases 1 to 305)		
JOURNAL	Bonaldo,M.F., Lennon,G. and Soares,M.B.		
MEDLINE	Normalization and subtraction: two approaches to facilitate gene		
COMMENT	discovery		
	Genome Res. 6 (9), 791-806 (1996)		
	On Jun 22, 1998 this sequence version replaced gi:3247553.		
	Contact: Chin, H		
	National Institute of Mental Health		
	6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD		
	20892-9643, USA		
	Tel.: 301 443 1706		
	Fax: 301 443 9890		
	Email: MEST@mail.nih.gov		
	The sequence contained an oligo-dT track that was present in the		
	oligonucleotide that was used to prime the synthesis of first		
	strand cDNA and therefore this may represent a bonafide poly A		
	tail. The sequence tag present in the cDNA between the NotI site		
	and the Oligo-dT track served to identify it as a clone from the		
	normalized basal ganglia library cDNA Library Preparation: M.B.		
	Soares Lab Clone distribution: NIH BMAP cDNA clones will be made		
	available by the means that is soon to be determined. When NIH		
	determines the means for distribution of the BMAP cDNA clones, this		
	record will be updated accordingly when that means is determined.		
	Seq primer: M13 Forward		
FEATURES	POLYA=yes.		
source	Location/Qualifiers		
	1..305		
	/organism="Mus musculus"		
	/strain="C57BL/6J"		
	/db_xref="taxon:10090"		
	/clone="UI-M-BH2.2-aos-g-02-0-UI"		
	/clone_id="NIH_BMAP_M.S3.2"		
	/dev_stage="27-32 days"		
	/lab_host="DHI0B (Life Technologies)"		
	/note="Vector: pRT3D-Pac (Pharmacia) with a modified		
	polylinker. Site.1: Not I; Site.2: Eco RI; The		
	NIH_BMAP_M.S3.2 library is a subtracted library of a		
	series, ultimately derived from a mixture of individually		
	tagged normalized libraries from ten regions of the mouse		
	brain (cerebellum, brain stems, olfactory bulbs,		
	hypothalamus, cortex, amygdala, basal ganglia, pineal		
	gland, striatum, hippocampus) after a series of		
	subtractions to reduce the representation of cDNAs from		
	which ESTs had already been generated. The following		
	serially subtracted libraries were generated in this		
	process: NIH_BMAP_M.S3.2, NIH_BMAP_M.S2, NIH_BMAP_M.S1.		
	The subtracted library (NIH_BMAP_M.S3.2) was constructed		
	as follows: PCR amplified cDNA inserts from NIH_BMAP_M.S2		
	clones from which 3' ESTs had been derived was used as a		
	driver in a hybridization with the NIH_BMAP_M.S2 library		
	in the form of single-stranded circles. The remaining		


```

/tissue_type="pooled"
/lab_host="DH10B"
/notice="vector: pCMV_SPORT6; site.1: xbaI; site.2: xhoI;
library made from pooled tissue from testis, ovary,
endometrium, hypochalamus, pituitary, and placenta."
BASE_COUNT 36 a 64 c 73 g 35 t
ORIGIN

```

Query Match	3.8%	Score 150.8;	DB 81;	Length 208;
Best Local Similarity	85.7%;	Pred. No. 4.6e-24;		
Matches 180;	Conservative 0;	Mismatches 27;	Indels 3;	Gaps 1;

2032 agccagcggggaacctccaacccggcgcccacacgcgttgctggggcgccgatgttggcttc 2091
||||| |||| | |||| | |||| | |||| | |||| | ||||
2 AGCCAGCGGGGCCCCCAAGCCGGGCTCACAC--AGTGGGGGCCGGACTGGCAGGCTTC 58

2092 ctctctggtgttctctgcagcatccctcactctctccgattgtrcgccgtrcagcagcgt 2151
 |||||
 59 TTCCTGGGGGTTCTTGCAGCATCCCTGACTCTTCTCCGATTTGGTCGGCGTCAGCAGCG 118

2152 cggcgacagaggagcttcagctagagacaaggtggtgttagactggtggtccaccct 2211
+ |||||
119 CAGCGACAGAGGACTTCTGGCTTAGAGACACAGCTGGGCTTGAGACTGGGGTCCACGCA 178
+ |||||

Y	2212	tcctggaccacaagctatagtcaggaccct	224
b	179	TCCGGACCAACAAGCTACAGCCAGACCCCT	208

	EST	11-NOV-1999
AU329323	257 bp mRNA	
AU329323	RIKEN full-length enriched, adult male medulla oligonucleotide	
Mus musculus CDNA clone 6330501J74.3'	similar to AB021291 Mus musculus mRNA for semaphorin W, mRNA sequence.	
AF050182		

CCESION	AV329323	GI:6369375
ERSION	AV329323.1	
EWK RDS	EST.	
OURC3	house mouse.	
ORGANISM	Mus musculus	

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 257)
Kono, H., Aizawa, K., Akahira, S., Akiyama, J., Enomoto, M., and
Akiyama, H.

YUKAWA, S., ISHIKAWA, T., HATA, A., HAYASHI, N., KODAZAKI, T., HOLL, R., ISHII, Y., ISHIKAWA, T., ITOH, M., IZAWA, M., NAKAO, K., KAGAOKA, I., KAI, C., KAWA, T., KIKUCHI, N., KOJIMA, Y., KOYA, S., KUSABEKE, M., MATSUYAMA, T., MIKI, R., MIZUNO, Y., NAKAMURA, M., ODA, H., OKAZAKI, Y., OWA, C., OZAWA, Y., SAITO, H., SANO, M., SATO, K., SHIBATA, K., SHIBATA, Y., SHIGEMOTO, Y., SHIRAKI, T., SOGAE, Y., SUGIHARA, Y., SUZUKI, H., SUZUKI, H., TAKASHI, F., TALENO, M., TOMIYAMA, N., TSUNODA, Y., WATSHIKI, A., WATANABE, S., YAMAMURA, T., YASUNISHI, A., YOKOTA, T., YOSHIKI, A., YOSHINO, M., MURAMATSU, M., and HAYASHIZAKI, Y. RIKEN Mouse Esch (Konno, H., et al.)

JOE RNAL
OMME VT

Genome Exploration Research Group, Life Science Tsukuba Center,
Genome Science Laboratory
The Institute of Physical and Chemical Research (RIKEN), Genomic
Sciences Center
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: +81-298-36-9013
Fax: +81-298-36-9098
Email: genome-res@rct.riken.go.jp
[URL: http://genome-res.rct.riken.go.jp/](http://genome-res.rct.riken.go.jp/)
Sasaki, N., Izawa, M., Watabiki, M., Ozawa, K., Tanaka, T., Yoneda, Y.,
Matsuura, S., Carninci, P., Muramatsu, M., Okazaki, Y. and
Hayashizaki, Y.
Transcriptional sequencing: A method for DNA sequencing using RNA
polymerase. *Proc. Natl. Acad. Sci. U.S.A.* 95 (7), 3455-3460 (1998).
Itoh, M., Kishinai, T., Akiyama, J., Shibata, K., Izawa, M., Kawai, S.

Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M., Okazaki, Y., and Hayashizaki, Y. Automated filtration-based high-throughput plasmid preparation system. *Genome Res.* 9 (5), 463-470 (1999)

Carninci, P. and Hayashizaki, Y. High-efficiency full-length cDNA cloning. *Methods Enzymol.* 303 19-44 (1999)

Please visit our web site (<http://genome.rtc.riken.go.jp>) for further details.

FEATURES
source

```

/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="6330501j24"
/clone_1b="RIKEN full-length enriched, adult male medulla oblongata"

```

```

/sex="male"
/tissue_type="medulla oblongata"
/dev_stage="adult"
)
host="human"

```

/lad_host1- Drive
/note="Site_1: Salt; Site_2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
project of Genome Exploration Research Group in Riken

Genomic Sciences Center and Genome Science Laboratory in RIKEN, Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was

primed with a primer [5'
GAGAGAGAGAGATCCACAGACTCTTTT TTTTTTTTNN 3'], cDNA was
prepared by using triethylamine thermo-activated reverse

transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot - 10.0 and subtraction to Rot - 100.0. Second strand cDNA was prepared with the primer adaptor of

aldona cDNA was prepared with the primer sequence 5'-GACGAGAGATTCGAGTAATTAATTAATCCCCCCCCCCCCCCCCCCCC-3'. cDNA was cloned into the XhoI and BamHI sites. Vector: a modified pBluescript KS(+) after bulk excision.

BASE COUNT	72 a	59 c	55 g	71 t
------------	------	------	------	------

[illegible]

Query Match	5.3%	Score 138.8	DB 12	Length 237
Best Local Similarity	79.5%	Pred. No. 2.0e-21		
Matches 202; Conservative	0	Mismatches 47	Indels 5	Gaps 3

Qy 3729 ggagacctcaaacccccagtcgtatgtcaaccttcacagtgaggatllctcgtctcgtcttc 3788

QY 3789 cttgacagcagcctgtbaactactcacgagtcctcctt-ggtttgagatcccggtgctt 3847

Db 64 TTCACACACACCCTGTGAATATATCAAGAGTCCCTTGGGTTTCGGAGTAACCAAGTGCTT 123

Qy 3848 tgaataggaactttgycgttgcacctaacctgacgaacttgatcgttca-ttgtaaagt 3906

Db: 124 TCGATAGGATCTTTGGCAAGTAGTCTAACCTTAGGACATGTGTTATTATTTATTTAAAGTG 183

Db 184 GGCATATACCTACTTCGAGGGTGTCTGCAAGCATCAACGAGAGCAACGATGAAACATATA 243
 350/ gggatataaccatccatcccaagggtgcgcgcgcaaggatcaaaagaggaacgatac aaatata 3503
 cy gggatataaccatccatcccaagggtgcgcgcgcaaggatcaaaagaggaacgatac aaatata 3503

Qy	3964	gcattaccacagc	3977
Db	244	GCATTACCCAGAGC	257

RESULT 15

LOCUS	DEFINITION	IMAGE	EST	07-JAN-1999
A1365650	A1365650	490 bp	mRNA	
	ap20N08.x1 Schiller oligodendroglioma Homo sapiens cDNA clone			
	IMAGE:1955967 3' similar to TR-062179 Q62179 SENAPHORIN C ; , mRNA			

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM n'cleic - nucleic search, using sw model

Run on: June 3, 2000, 01:06:58 ; Search time 5907.01 Seconds
(without alignments)
134.490 Million cell updates/sec

Title: US-09-284-180-7

Perfect score: 196
Sequence: 1 aaattgtaccacagctgct.....ggtccacagacatagctcag 196

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Search: 4857316 seqs, 2026611650 residues

Total number of hits satisfying chosen parameters: 9714632

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_est1:*
2: em_est2:*
3: em_est3:*
4: em_est4:*
5: em_est5:*
6: em_est6:*
7: em_est7:*
8: em_est8:*
9: em_est9:*
10: em_est10:*
11: em_est11:*
12: em_est12:*
13: em_est13:*
14: em_est14:*
15: em_est15:*
16: em_est16:*
17: em_est17:*
18: em_est18:*
19: em_est19:*
20: gb_est1:*
21: gb_est2:*
22: gb_est3:*
23: gb_est4:*
24: gb_est5:*
25: gb_est6:*
26: gb_est7:*
27: gb_est8:*
28: gb_est9:*
29: gb_est10:*
30: gb_est11:*
31: gb_est12:*
32: gb_est13:*
33: gb_est14:*
34: gb_est15:*
35: gb_est16:*
36: gb_est17:*
37: gb_est18:*
38: gb_est19:*
39: gb_est20:*
40: gb_est21:*
41: gb_est22:*
42: gb_est23:*
43: gb_est24:*

44: gb_est25:*
45: gb_est26:*
46: gb_est27:*
47: gb_est28:*
48: gb_est29:*
49: em_est20:*
50: em_est21:*
51: em_est22:*
52: gb_est30:*
53: gb_est31:*
54: gb_est32:*
55: em_est23:*
56: em_est24:*
57: em_est25:*
58: em_est26:*
59: gb_est33:*
60: gb_est34:*
61: gb_est35:*
62: gb_est36:*
63: gb_est37:*
64: gb_est38:*
65: em_est27:*
66: em_est28:*
67: em_est29:*
68: em_est30:*
69: gb_est39:*
70: gb_est40:*
71: gb_est41:*
72: gb_est42:*
73: gb_est43:*
74: gb_est44:*
75: em_est31:*
76: em_est32:*
77: em_est33:*
78: em_est34:*
79: gb_est45:*
80: gb_est46:*
81: gb_est47:*
82: gb_gss1:*
83: gb_gss2:*
84: gb_gss3:*
85: gb_gss4:*
86: em_gss1:*
87: em_gss2:*
88: em_gss3:*
89: em_gss4:*
90: gb_gss5:*
91: gb_gss6:*
92: em_gss5:*
93: em_gss6:*
94: gb_gss7:*
95: gb_gss8:*
96: gb_gss9:*
97: em_gss7:*
98: em_gss8:*
99: gb_gss11:*
100: gb_gss10:*
101: em_gss9:*
102: em_gss10:*
103: em_gss11:*
104: em_gss12:*
105: gb_gss12:*
106: gb_gss13:*
107: gb_gss14:*
108: gb_gss15:*
109: gb_gss16:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES


```

COMM NT      Contact: Adams, MD
              The Institute for Genomic Research
              932 Clopper Road, Gaithersburg, MD 20878
              Tel: 3018699056
              Fax: 3018699423
              Email: mdadams@tigr.org
              Seq primer: M13 Reverse
              Location/Qualifiers
              Source
                1..364
                /organism="Homo sapiens"
                /db_xref="AFCC (Inhost):85555"
                /db_xref="taxon:9606"
                /clone="HIB071"
                /clone_lib="Infant Brain, Bento Soares"

BASE COUNT   70 a 100 c 106 g 85 t 3 others
ORIG N

Query Match  99.5%: Score 195; DB 20; Length 364;
Best Local Similarity 99.5%: Pred. No. 4.6e-49;
Matches 195; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 aaattgacacagctgtgctcctgtgtgctccgctactgaggtgacacagtgatata 60
    |||||||
Db 1 AAATGTACACAGCTGCTCCTGTGCTGCCCTCCCTACTGAGGTGACACAGTGAATATA 60
    |||||||

Oy 61 accaactgtgcccgtctccagagctgctcagagtgatcctctggcccaagaccagtcgt 120
    |||||||
Db 61 ACNACTGTGGCGCTCTCCAGAGCTCTCAGAGTGCATCTGTGGCCAGACCAAGTCTGT 120
    |||||||

Oy 121 gctctgagcttcctgctgctgagtggtgtgtgcccacatgctggagagcagtggtgtgc 180
    |||||||
Db 121 GCCTGAGCTTCGCGCTGATGATGTGTGTGCCCATGCGGAGACCGAGAGGTTGTGC 180
    |||||||

Oy 181 caagacatagctcag 196
    |||||||
Db 181 CAAGACATGAGTCAAG 196
    |||||||

RESULT 3
LOCUS       242773      212 bp      mRNA      EST      11-NOV-1994
DEFINITION  HSC07E021 normalized infant brain cDNA Homo sapiens cDNA clone
              c-0le02, mRNA sequence.
ACCESSION   242773
VERSIONS    242773.1 GI:570345
KEYWORDS    EST.
SOURCE      human.
            OR ANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Homiidae; Homo.
            1 (bases 1 to 212)
            Aufferay,C., Behar,G., Bols,F., Bouchard,C., da Silva,C.,
            Devignes,M.D., Duprat,S., Houlgatte,R., Juneau,M.N., Lamy,B.,
            Lorenz,F., Mitchell,H., Mariage-samson,R., Pletu,G., Pouillot,Y.,
            Sebastien-Kabatchnik,C. and Tessier,A.
            IMAGE: molecular integration of the analysis of the human genome
            and its expression
            C. R. Acad. Sci. III, Sci. Vie 318 (2), 263-272 (1995)
            95277534
            Contact: Genethon
            Genexpress-Genethon
            Genethon Centre de recherche sur le Genome Humain
            1,rue de l'Internationale, BP60 91002 EVRY Cedex, FRANCE
            Tel: 33169472800
            Fax: 33160778698
            Email: genexpress@genethon.fr
            Single read.
            Genexpress_library_id: C; Genexpress_sequence_id: y1c-0le02
            Seq primer: (-21)M13 universal.
            Location/Qualifiers
            1..212
            /organism="Homo sapiens"

COMM NT      Contact: Adams, MD
              The Institute for Genomic Research
              932 Clopper Road, Gaithersburg, MD 20878
              Tel: 3018699056
              Fax: 3018699423
              Email: mdadams@tigr.org
              Seq primer: M13 Reverse
              Location/Qualifiers
              Source
                1..364
                /organism="Homo sapiens"
                /db_xref="AFCC (Inhost):85555"
                /db_xref="taxon:9606"
                /clone="HIB071"
                /clone_lib="Infant Brain, Bento Soares"

BASE COUNT   70 a 100 c 106 g 85 t 3 others
ORIG N

Query Match  99.5%: Score 195; DB 20; Length 364;
Best Local Similarity 99.5%: Pred. No. 4.6e-49;
Matches 195; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 aaattgacacagctgtgctcctgtgtgctccgctactgaggtgacacagtgatata 60
    |||||||
Db 1 AAATGTACACAGCTGCTCCTGTGCTGCCCTCCCTACTGAGGTGACACAGTGAATATA 60
    |||||||

Oy 61 accaactgtgcccgtctccagagctgctcagagtgatcctctggcccaagaccagtcgt 120
    |||||||
Db 61 ACNACTGTGGCGCTCTCCAGAGCTCTCAGAGTGCATCTGTGGCCAGACCAAGTCTGT 120
    |||||||

Oy 121 gctctgagcttcctgctgctgagtggtgtgtgcccacatgctggagagcagtggtgtgc 180
    |||||||
Db 121 GCCTGAGCTTCGCGCTGATGATGTGTGTGCCCATGCGGAGACCGAGAGGTTGTGC 180
    |||||||

Oy 181 caagacatagctcag 196
    |||||||
Db 181 CAAGACATGAGTCAAG 196
    |||||||

RESULT 4
LOCUS       AW258432      539 bp      mRNA      EST      23-DEC-1999
DEFINITION  uc34h08.y1 NCI CGAP Mam5 mus musculus cDNA clone IMAGE: 5' similar
              to TR:064151 Q64151 SEMAPHORIN I ;, mRNA sequence.
ACCESSION   AW258432
VERSION     AW258432.1 GI:6631413
KEYWORDS    EST.
SOURCE      house mouse.
            ORGANISM Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
            1 (bases 1 to 539)
            NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
            National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
            Unpublished (1997)
            On May 18, 1998 this sequence version replaced gi:3137576.
            Other ESTs: uc34h08.x1
            Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert.Strausberg@nih.gov
            Tissue procurement: Lothar Hennighausen Ph.D., Robin Humphreys
            cDNA library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
            DNA sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LNL at:
            www.bio.lnl.gov/db/ftp/image/image.html

FEATURES             Seq primer: -40RP from Glbco
                    High quality sequence stop: 412.
                    Location/Qualifiers
                    1..539
                    /organism="Mus musculus"
                    /strain="C57/B6"
                    /db_xref="taxon:10090"
                    /clone="IMAGE:"
                    /clone_lib="NCI CGAP Mam5"
                    /tissue_type="tumor, gross tissue"
                    /dev_stage="7 months"
                    /lab_host="DH10B"

```

BASE COUNT	106 a	157 c	141 g	135 t
ORIGIN	/note="Organ: mammary; Vector: PCWV-SPORT6; Site:1: SalI Site:2: NotI; Cloned unidirectionally. Primer: Oligo dr library constructed by Life Technologies, investigators providing samples: Lothar Hennigshausen/Robin Humphreys, NIH"			

Query Match	23.5%	Score 46	DB 79	Length 539
Best Local Similarity	54.0%	Pred. No. 0.00056		
Matches 94	Conservative 0	Mismatches 80	Indels 0	Gaps 0

QY 18 gctccgtgtgctccgtaactgagtgacacaaagtgaatacaccaactgtggcgcgtct 77

Db 32 GCCTTTCCTGGCTCCCGCTCTCAGCTGGTTACAGCTGTCTCTGGCCGACTGCACAAAGTA 91

[illegible]

138 acataadatatgcccataccgagaaacacgaacattatccaagacataga 191

Db 152 CAGCCGCTGTGGCCACCAAGTGGTCGCTCGGGGTCCTTCTGTGTCACACA 205

11

AW345392	257 bp	MRNA	01-FEB-2000
LOCUS			
AW345392			

DEFINITION 24958 MARC 4BOV BOS CAIUS CDNA 5' , mRNA sequence
ACCESSION AM345392

KEYWC.RDS EST.
BOUWPC BOG taurus

ORC:ANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

Eurhinella, Ceraclinoactylia, Xanthanella, Felcia, Bovinea, Bovinae; Bos.

AUTHORS
Smith, T. P. L., Casas, E., Stone, R. T., Heaton, M. P., Grosse, W. M.,
Bennett, G. A., Fahrenkrug, S. C., Freking, B. A., Rohrer, G. A. and
Bennett, G. A.

TYPE Design and use of four pooled tissue normalized cDNA libraries for

Unpublished (2000)
On Tue 22 1000 this resource moved to 30/6005

Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center

PO Box 100, Clay Center, NE 68533-V100, USA
Tel: 402 762 4366

Email: smith@email.marc.usda.gov
Single base sequencing Bases called and trimmed with phred

v0.980904.e. Vector identified by cross_match with the -minscore 20 and -mismatch 12 options.

FORN FIELDS
 FORWARD: AGGAACAGCTATGACCAT

Plate: 13 row: L column: 1
Seq primer: ATTTAGGTGACCTATAG

FEATURES	Location/Qualifiers
SOURCE	1. .257

```
/organism= "Drosophila"
/db_xref="taxon:9913"
```

```
/tissue_type="pooled"
/lab host="DH10B"
```

```

/note="vector; pcmv spkto; site_1: abd; site_2: ano;
Library made from pooled tissue from day 20 and day 40

```

BASE COUNT	50 a	80 c	82 g	45 t
00000	00000	00000	00000	00000
00001	00001	00001	00001	00001
00002	00002	00002	00002	00002
00003	00003	00003	00003	00003
00004	00004	00004	00004	00004
00005	00005	00005	00005	00005
00006	00006	00006	00006	00006
00007	00007	00007	00007	00007
00008	00008	00008	00008	00008
00009	00009	00009	00009	00009
00010	00010	00010	00010	00010
00011	00011	00011	00011	00011
00012	00012	00012	00012	00012
00013	00013	00013	00013	00013
00014	00014	00014	00014	00014
00015	00015	00015	00015	00015
00016	00016	00016	00016	00016
00017	00017	00017	00017	00017
00018	00018	00018	00018	00018
00019	00019	00019	00019	00019
00020	00020	00020	00020	00020
00021	00021	00021	00021	00021
00022	00022	00022	00022	00022
00023	00023	00023	00023	00023
00024	00024	00024	00024	00024
00025	00025	00025	00025	00025
00026	00026	00026	00026	00026
00027	00027	00027	00027	00027
00028	00028	00028	00028	00028
00029	00029	00029	00029	00029
00030	00030	00030	00030	00030
00031	00031	00031	00031	00031
00032	00032	00032	00032	00032
00033	00033	00033	00033	00033
00034	00034	00034	00034	00034
00035	00035	00035	00035	00035
00036	00036	00036	00036	00036
00037	00037	00037	00037	00037
00038	00038	00038	00038	00038
00039	00039	00039	00039	00039
00040	00040	00040	00040	00040
00041	00041	00041	00041	00041
00042	00042	00042	00042	00042
00043	00043	00043	00043	00043
00044	00044	00044	00044	00044
00045	00045	00045	00045	00045
00046	00046	00046	00046	00046
00047	00047	00047	00047	00047
00048	00048	00048	00048	00048
00049	00049	00049	00049	00049
00050	00050	00050	00050	00050
00051	00051	00051	00051	00051
00052	00052	00052	00052	00052
00053	00053	00053	00053	00053
00054	00054	00054	00054	00054
00055	00055	00055	00055	00055
00056	00056	00056	00056	00056
00057	00057	00057	00057	00057
00058	00058	00058	00058	00058
00059	00059	00059	00059	00059
00060	00060	00060	00060	00060
00061	00061	00061	00061	00061
00062	00062	00062	00062	00062
00063	00063	00063	00063	00063
00064	00064	00064	00064	00064

Query Match	23.0%	Score	45	DB	80	Length	257
Best Local Similarity	67.7%	Pred. No.	0.00085				
Matches	63	Conservative	0	Mismatches	30	Indels	0
						Gaps	0

Qy 62 ccacacgctgccccgctctccacagagctctgctcagagctgcatctctgccccagacccagctctg 121
Db 6 ccacatgcagcctgtatccacagagctgcagggactgcctctgcgccccggagacccctactacgc 65

Db 66 CTTGAGCGGCTGGAGGTGTGCTTCCCTCAGCC 98

RESULT	6
A1787249/c	
LOCUS	
DEFINITION	
A1787249	398 bp
AF58910 v1	Susaco mouse
	liver m[1]e Mus musculus
	CNN2 clone
	02-JUL-1999

ACCESSION	SEQUENCE	CO IN.	COZ	SEMPHROLIN C	UNITARY
AI787249.1	GI:5334965	1924122	3	SMITHI	
AI787249					

SOURCE	
ORGANISM	house mouse.
	<i>Mus musculus</i> .
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

REFERENCE
1 (bases 1 to 398)
Marta, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Papad, D., Harvey, N., Schuk, R.

Waterston, R. and Wilson, R.
The WashU-NCI Mouse EST Project 1999
Unpublished (1999)

Other E-Strs: u358a10.y1
Contact: Marra M/Washu-NCI Mouse EST Project 1999
Washington University School of Medicine
444 E. Forest Park Parkway, Box 8501 St. Louis MO 63108 USA

Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through LINT; contact the

MGI:980414
Possible reversed clone: similarity on wrong strand
Seq. primer: custom primer used

```
FEATURES
source
Location/Qualifiers
1. .398
/organism="Mus musculus"
/strain="C57Bl"
```

```

__clone__IMAGE:1924122"
/clone_1lb"Sugano mouse liver m1a"
/sex="female"

```

```

/lab_host="DH10B"
/note="Organ: liver"
(CACCTGCTG); Site_1: DraIII
(CACCTGCTG); Site_2: DraIII (CACCATCTG); 1st strand cDNA
(CACCTGCTG); Site_3: DraIII

```

ligated to a DraIII adaptor (TGTGGCCCTTACCTGG), digested and cloned into distinct DraIII sites of the pME185-FL3 vector (5' site CACTGTGTG, 3' site CACCACTGTG). XhoI should

performed to exclude fragments <1.5kb. Library constructed by Dr. Sumio Sugano (University of Tokyo Institute of Medical Science). Custom primers for

BASE COUNT					
92	a	104	c	118	g
				83	t
					1 others

Oy	18	gaccctgtgtggctcccgacacgaggtgacacaagtacacccaactgtggccgct	77
Db	542	gctcttttggccggctccctcttcagctgtgtgacgctgcccgtggccgcatgtcattaaagta	483
Oy	78	ccagagctgctacagagtgatcctctggccgaggaaccctcgtctgcccgtgagcttcgcgt	137
Db	482	tgcgttcctctgtgacagactgtgtctctgccccgggahcccttattggccctggagcgtcaacac	423
Oy	138	gagatgagtgtgtgcccattcccg	161
Db	422	cagccgctgtgtgcccgtggctgg	399

RESULT	9
A1638381.C	
LOCUS	
DEFINITION	710 bp mRNA EST 14-DEC-1999
ACCESSION	A1638881
VERSIONS	t08d08.x1 NC1 CGAP GC6 Homo sapiens cDNA clone IMAGE:2240175 3
KEYWORDS	similar to TRIQ64151 Q64151 SEMAPHORIN I ; , mRNA sequence.
SOURCE	A1638881.1 GI:4691115
EST.	
HUMAN.	

REF:ENCE
AUT:ORS
TIT:JE
JOL:RNAL
COMM:VT

1 (bases 1 to 710)
NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CCAP).
Tumor Gene Index
Unpublished (1997)
On Mar 10, 1998 this sequence version replaced gi:2948370.
Contact: Robert Streuberg, Ph.D.

Unpublished (1998) this sequence version replaced g1:2948370.
On Mar 10, 1998 this sequence version replaced g1:2948370.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Christopher A. Moskalko, M.D., Ph.D., Michael
R. Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldi, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-60GAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bdrp/image/image.html

Insert Length: 793 Std Error: 0.00
Seq primer: -40UP from GIBCO
High quality sequence stop: 444.

FEATLRES	Location/Qualifiers
source	1. .710

BASE COUNT	171 a	190 c	183 g	165 t	1 others
ORIGIN					

```

Query Match      20.4%; Score 40; DB 52; Length 710;
Best Local Similarity 54.9%; Pred. No. 0.04;
Matches 79; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

```

QY	18	ggcccttggtggtctccgctactgaagtgacacaaagtgaatacaaccaactgtgcccgtc	77
Db	543	gctctttggccgctcccgcttccactgtgtcagctgcccggccgcatgcatttaagta	484
QY	78	ccagagctgtcagaagtgcattcctctgcccagagaccagctctgtgcttgagcttcggt	137
Db	483	tggcttccctgtgacgactgtgtcttgcggccggagaccctatttcgcttgagcgtaacac	424
QY	138	ggatgagtgctgtgcccattgccg	161
Db	423	cagccgctgtgtgcccgtggctgg	400

RESULT	10
T08621	
LOCUS	T08621
DEFINITION	T08621 206 bp mRNA EST T08621 Infant Brain, Bento Soares Homo sapiens CDNA clone HIBBH20 5' end, mRNA sequence.
ACCESSION	T08621
VERSION	T08621.1 GI:389649
KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens <i>Eutheria; Amniota; Chordata; Craniata; Vertebrata; Mammalia;</i>

REFERENCE
1 (Cases 1 to 206)
AUTHORS Adams,M.D., Soares,M.B., Kerlavage,A.R., Fields,C. and Venter,J.C.
TITLE Rapid cDNA sequencing (expressed sequence tags) from a
directionally cloned human infant brain cDNA library
JOURNAL Nature Genet. 4, 373-380 (1993)
MEDLINE 94004965
COMMENT Contact: Adams, MD
The Institute for Genomic Research

COMMENT

Contact: Adams, MD
The Institute for Genomic Research
933 Clopper Road, Gaithersburg, MD 20878
Tel.: 301.859.0056
Fax: 301.859.4423
Email: moadams@tigr.org
Seq primer: MJ Reverse.

FEATURES	Location/Qualifiers
source	1. .206

```

/clone_1lb="Infant Brain, Bento Scores"
BASE COUNT      30 a      66 c      56 g      50 t      4 others
ORIGIN

```

Query Match	19.2%	Score	37.6	DB	20	Length	206
Best Local Similarity	52.8%	Pred. NO.	0.13				
Matches	76	Conservative	0	Mismatches	68	Indels	0
						Gaps	0

Qy 18 gcccttggttgccgcgcctactgaggtgacacaaatgaataacacaaatcgtggccgtct 77
 Db 6 gctctttttcccgactcccgcnctcagctgctgacactctccgtggccgacactgatgaagta 65
 Qy 78 ccagagcgcgtcaagatgcatccctggccccaagaccacagtcgtgacctggagatccgct 137
 Db 66 tgcctcctcgtgcagactgtgtctctggcccggnaccctattgcctgtgagacctcaaac 125
 Qy 138 ggaatgagtgctggtggcccatgcccgg 161
 Db 126 cagccgcntgintggccgtggctgg 149

RESULT	11			
M78717				
LOCUS	M78717	265 bp	EST	26-MAY-1992
DEFINITION	E890865 Hippocampus, Strata gene (cat. #36205) Homo sapiens cDNA.			
ACCESSION	Clone H8C8D08, mRNA sequence.			
VERSION	M78717			
	M78717.1	GI:273032		

```
KEYW RDS EST.  
SOUR E human.  
OR ANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
REF ENCE 1 (bases 1 to 265)  
AU HORS Adams,M.D., Dubnick,M., Kerlavage,A.R., Moreno,R., Kelley,J.M.,  
Utterback,T.R., Nagle,J.W., Fields,C. and Venter,J.C.  
TI LE Sequence identification of 2,375 human brain genes  
JOL RNAL Nature 355, 632-634 (1992)  
ME:LIN 9216812  
COMM NT Contact: Kerlavage, AR  
Bioinformatics  
The Institute for Genomic Research  
9712 Medical Center Drive, Rockville, MD 20850 USA  
Tel.: 3018699056  
Fax: 3018699423  
Email: arkerlav@tigr.org  
Seq primer: M13 Forward.
```

Location/Qualifiers

```
1..265  
/organism="Homo sapiens"  
/db_xref="ATCC (inhost)":78133"  
/db_xref="GDB:DOSI705E"  
/db_xref="taxon:9606"  
/clone="HHCMD08"  
/clone.lib="Hipocampus, Striatum (cat.#36205)"  
/note="Vector: lambdaZAP-II; Female, 2 years; oligo-dT +  
random primed cDNA synthesis; lambdaZAP-II vector, 1.0kb  
average insert size."  
BASE COUNT 45 a 86 c 78 g 50 t 6 others  
ORIG N
```

Ovary Match 18.7% Score 36.6 DB 20 Length 265;

Bet Local Similarity 56.2%; Pred.No.0.3; Indels 0 Gaps 0

Mach 63 Conservative 1; Mismatches 48;

Dy 18 gccctcgttgcccccagtcagtgcacacgtaaacaccactgtggcgctc 77
||| | |||| | | |||| | | |||| | | |||| | | |||| | |
Db 36 gcctgatgccggcctcanantcgaggctaattccagngccccatggccaactgcagcctga 95

Dy 78 ccagaagcgtcacagtagtatccctgcgcccagaagaccagtcgttgtcctgagc 129
| |||||| | |||| |||| |||| |||| |||| |||| |||| ||||
Db 96 CAGGAACCTGTGGCACTCCTCGGCCGGAACCCTACTGTGCYTKKAGC 147

```
RESU T 12  
LOCUT AM44 785 mRNA EST 17-FEB-2000  
DEFINITIO 89922 MARC IBOV Bos taureus CDNA 5' , mRNA sequence.  
ACCESSION AM447785  
VERS ON AM447785.1 GI:6989572  
KEYW RDS EST.  
SOUR E Bos taureus.  
OR ANISM Bos taureus.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
Euheria; Cetartiodactyla; Rumiantia; Pecora; Bovidae; Bovidae;  
Bovinae; Bos.  
1 (bases 1 to 488)  
Smith.T.P.L., Casas,E., Stone,R.T., Heaton,M.P., Grose,W.M.,  
Benett,G.A., Fehrenkrug,S.C., Fraking,B.A., Rohrer,G.A. and  
Keeler,J.W.  
Design and use of four pooled tissue normalized cdna libraries for  
EST discovery in cattle  
Unpublished (2000)  
On Jan 6, 2000 this sequence version replaced gi:667495.  
Contact: Smith TP,  
USDA, ARS, US Meat Animal Research Center  
PO Box 166 Clay Center, NE 68933-0166, USA  
Tel: 402 762 4366  
Fax: 402 762 4390  
Email: smithemail.marc.usda.gov
```

FEATURES	Source	Location/Qualifiers
	1..488	
	/organism="Bos taurus"	
	/db_xref="taxon:9913"	
	/clone_id="MARC 1BOY"	
	/tissue_type="pooled"	
	/lab_host="DH10B"	
	/note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI; Library made from pooled tissue from lymph node, ovary, fat, hypothalamus, and pituitary."	
BASE COUNT	65 a	187 c 142 g 94 t
ORIGIN		
Query Match	18.7%	Score 36.6; DB 81; Length 488;
Best Local Similarity	55.0%	Pred. No. 0.37;
Matches	72; Conservative	0; Mismatches 59; Indels 0; Gaps 0;
Oy	24	ggttgctcccgctactggtgagtgacacaaagtgaatcaccaactgtggcgcgtccagag 83
Db	1	GGCCGGTTCACACTGACGCGGTGTGTGACAGCCCTGTGGCTCTTCTCGGGAAGCAATGCGAC 60
Oy	84	ctgtctcagaagtgcacatcctgtgcccagagaccagctctgtgctctggagcttcgcgtgatga 143
Db	61	CTGCGAGGACGTGCGGTGCTTCCGCGGAGCCCTACTGCGCTGTGAGCCCGCGCGTGGCGCG 120
Oy	144	gtgtgtgtggccc 154
Db	121	CTGTGTGCGCC 131
RESULT 13		
LOCUS	AM437778	365 bp mRNA EST 14-FEB-2000
DEFINITION	79230 MARC 1BOV Bos taurus cDNA 5', mRNA sequence.	
ACCESSION	AM437778	
VERSION	AM437778.1	GI:6973084
KEYWORDS	EST.	
SOURCE	Bos taurus.	
ORGANISM	Bos taurus	
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;	
	Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea; Bovidae;	
	Bovinae; Bos.	
	1 (bases 1 to 365)	
REFERENCE	Smith,T.P.L., Casas,E., Stone,R.T., Heaton,M.P., Grosse,W.M.,	
AUTHORS	Bennett,G.A., Fahrenkrug,S.C., Freking,B.A., Rohrer,G.A. and	
	Keefe,J.W.	
TITLE	Design and use of four pooled tissue normalized cDNA libraries for	
	EST discovery in cattle	
JOURNAL	Unpublished (2000)	
COMMENT	On Jan 6, 2000 this sequence version replaced gi:6675875.	
	Contact: Smith TPL	
	USDA, ARS, US Meat Animal Research Center	
	PO Box 166, Clay Center, NE 68933-0166, USA	
	Tel: 402 762 4366	
	Fax: 402 762 4390	
	Email: smtlhemail.marc.usda.gov	
	Single pass sequencing. Bases called and trimmed with phred	
	v0.980904.e. Vector identified by cross_match with the -minscore 20	
	and -minmatch 12 options.	
	PCR Primers	
	FORWARD: AGGAACAGCTATGACCAT	
	BACKWARD: GTTTTCCAGTCCAGCAG	
	Plate: 76 row: L column: 1	
	Seq primer: ATTAGGAGACACTATAG.	

```

FEAT1 RES      Location/Qualifiers
SOURCE          1..365
                /organism="Bos taurus"
                /db_xref="taxon:9913"
                /clone_1lb="MARC 1BOV"
                /library_type="pooled"
                /lab_host="DH10B"
                /note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;
                library made from pooled tissue from lymph node, ovary,
                fat, hypothalamus, and pituitary."
BASE COUNT     49 a 138 c 104 g 74 t
ORIGIN
Que cy Match   18.3%; Score 35.8; DB 81; Length 365;
Bec Local Similarity 55.1%; Pred. No. 0.58;
Matches 70; Conservative 0; Mismatches 57; Indels 0; Gaps 0;
Oy 28 ggtccgactgaggtgacacaaagtaatacaacaaactgtgagcgtctccagagctgc 87
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 56 ggttccaactcagcggtggtgcagcgcctctgaccttcttcggcgaacatggcacttgc 115
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 88 tcaggtgacatccctgcccagagaccagctctctgtgcttgagcttcaggctgagataggt 147
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 116 gagcagctgctgtcttcccgagacccttactgctgacctgagaccgccgcctgacctctgt 175
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 148 gtgagcc 154
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 176 gtgcgcc 182
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 14
AAC87573      487 bp      mRNA      EST      13-AUG-1997
LOCUS        z55297.1 NC1.CGAP.GC81 Homo sapiens CDNA clone IMAGE:701148 5'
DEFINITION   similar to FR:G886813 G886813 COLLAPSIIN-4 ;, mRNA sequence.
ACCESSION   AA287573
VERSION      AA287573.1 GI:1933255
KEYWORDS     EST.
SOURCE       human.
ORGANISM     Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1 (bases 1 to 487)
            NC1.CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
            National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
            Unpublished (1997)
            On Sep 12, 1996 this sequence version replaced gi:1288732.
            Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            This clone is available royalty-free through LLNL; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -28ml3 rev2 ET from Amersham
            High quality sequence stop: 474.
            Location/Qualifiers
                1..487
                /organism="Homo sapiens"
                /db_xref="GDB:5740583"
                /db_xref="taxon:9606"
                /clone_image="IMAGE:701148"
                /clone_1lb="NC1.CGAP.GC81"
                /library_type="germline center B cell"
                /lab_host="DH10B"
                /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                was prepared from human tonsillar cells enriched for
                germinal center B cells by flow sorting (CD20+, IgD-),
                provided by Dr. Louis M. Staudt (NCI), Dr. David Allman
                (NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was
                primed with a Not I - oligo(dT) primer
                15'-GTGTACCAATCTAGAGGAGCGCCGCTCATTTTTTTTTTTTTTTTTT-

```

[illegible]

	Matches	74:	Conservative	0:	Mismatches	46:	Indels	3:	Gaps	1:
Oy	32	ccccctactgaggtgacacaagtgaatacaacacactgtg--gcgctctccagagctgct	88							
Db	691	CCCATGTTAAGGTGAGACAATTAATTCAAAGAACAATGTTTCCGCTGCAAAAGGCATT	632							
Oy	89	cagagtcgcatcctgagccagaccagctctgtgctcctgagagctccggcttgatgagtgtg	148							
Db	631	CAGGCTGTGGCCTGGCCCAAAACCATTTCTTTATATAGCTTATGGCCCTGTGATGGAG	572							
Oy	149	tgg 151								
Db	571	GGG 569								

Search completed: June 3, 2000, 01:07:02
 Job time: 58233 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM n cleic - nucleic search, using sw model

Run : June 3, 2000, 01:06:50 ; Search time 5907.01 Seconds
(Without alignments)
1208.348 Million cell updates/sec

Title : US-09-284-180-5

Perf ct score : 1761
Sequence : 1 ggggggtgcctctatgctgc.....catgtatgaacatccatc 1761

Scoring table : IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched : 4857316 seqs, 2026611650 residues

Total number of hits satisfying chosen parameters : 9714632

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:
1: em_est1:*
2: em_est2:*
3: em_est3:*
4: em_est4:*
5: em_est5:*
6: em_est6:*
7: em_est7:*
8: em_est8:*
9: em_est9:*
10: em_est10:*
11: em_est11:*
12: em_est12:*
13: em_est13:*
14: em_est14:*
15: em_est15:*
16: em_est16:*
17: em_est17:*
18: em_est18:*
19: em_est19:*
20: gb_est1:*
21: gb_est2:*
22: gb_est3:*
23: gb_est4:*
24: gb_est5:*
25: gb_est6:*
26: gb_est7:*
27: gb_est8:*
28: gb_est9:*
29: gb_est10:*
30: gb_est11:*
31: gb_est12:*
32: gb_est13:*
33: gb_est14:*
34: gb_est15:*
35: gb_est16:*
36: gb_est17:*
37: gb_est18:*
38: gb_est19:*
39: gb_est20:*
40: gb_est21:*
41: gb_est22:*
42: gb_est23:*
43: gb_est24:*

44: gb_est25:*
45: gb_est26:*
46: gb_est27:*
47: gb_est28:*
48: gb_est29:*
49: gb_est20:*
50: gb_est21:*
51: gb_est22:*
52: gb_est30:*
53: gb_est31:*
54: gb_est32:*
55: gb_est23:*
56: gb_est24:*
57: gb_est25:*
58: gb_est26:*
59: gb_est33:*
60: gb_est34:*
61: gb_est35:*
62: gb_est36:*
63: gb_est37:*
64: gb_est38:*
65: gb_est27:*
66: gb_est28:*
67: gb_est29:*
68: gb_est30:*
69: gb_est39:*
70: gb_est40:*
71: gb_est41:*
72: gb_est42:*
73: gb_est43:*
74: gb_est44:*
75: gb_est31:*
76: gb_est32:*
77: gb_est33:*
78: gb_est34:*
79: gb_est45:*
80: gb_est46:*
81: gb_est47:*
82: gb_est48:*
83: gb_est49:*
84: gb_est50:*
85: gb_est51:*
86: gb_est52:*
87: gb_est53:*
88: gb_est54:*
89: gb_est55:*
90: gb_est56:*
91: gb_est57:*
92: gb_est58:*
93: gb_est59:*
94: gb_est60:*
95: gb_est61:*
96: gb_est62:*
97: gb_est63:*
98: gb_est64:*
99: gb_est65:*
100: gb_est66:*
101: gb_est67:*
102: gb_est68:*
103: gb_est69:*
104: gb_est70:*
105: gb_est71:*
106: gb_est72:*
107: gb_est73:*
108: gb_est74:*
109: gb_est75:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result NC	Score	Query Match	Length	DB	ID	Description
1	476.4	27.1	503	34	AA459837	AA459837 zx50g12.r
2	356.2	20.2	364	20	T09073	T09073 EST06966.in
3	349.2	19.8	517	22	R54387	R54387 Y878f12.r1
4	338.4	19.0	337	20	Z45329	Z45329 HSC2M6091.n
5	328.4	18.6	437	23	H24181	H24181 ym55f02.r1
6	326.4	18.5	347	22	H10623	H10623 ym08f07.r1
7	308.2	17.5	376	81	AW436751	AW436751 77209.MAR
8	229.4	13.0	592	106	AQ352219	AQ352219 CTBI-E1-
9	167	9.5	208	81	AW436702	AW436702 77141.MAR
10	141	8.0	212	20	Z42773	Z42773 HSC0E021.n
11	112.6	6.4	232	80	AW346693	AW346693 29273.MAR
12	112.6	6.4	232	80	AW346698	AW346698 29282.MAR
13	98.4	5.6	490	40	AW365650	AW365650 ap20h08.x
14	93.2	5.3	619	64	AL118624	AL118624 DKF2P61F
15	84.6	4.8	442	48	AW55178	AW55178 UT-R-C2P
16	76.2	4.3	638	79	AW245910	AW245910 2832888.5
17	72.4	4.1	942	69	AW078986	AW078986 AW078986
18	67.8	3.9	469	42	AW141990	AW141990 ooi7910.x
19	67.4	3.8	555	52	AW1641288	AW1641288 fci3a11.y
20	57.8	3.3	925	82	CNS0091P	AL053013 Drosoph11
21	55.6	3.2	398	60	AW1787249	AW1787249 u158a10.x
22	50	2.8	203	31	AA326134	AA326134 EST29247
23	46.8	2.7	539	79	AW258432	AW258432 ug34h08.y
24	46.6	2.6	492	27	AA023538	AA023538 mh75d07.t
25	45.8	2.6	498	29	AA190645	AA190645 zq44a06.x
26	44	2.6	257	80	AW345392	AW345392 24958.MAR
27	44	2.5	900	54	C82295	C82295 C82295 Leuk
28	43.6	2.5	320	74	AW227926	AW227926 up19b05.y
29	43.6	2.5	889	74	AW050267	AW050267 AW050267
30	43.4	2.5	624	74	AW328571	AW328571 ds03d12.x
31	42.6	2.4	997	82	CNS006DN	AL065132 Drosoph11
32	41.8	2.4	1049	83	CNS015B1	AL105223 Drosoph11
33	41.8	2.4	266	20	Z28925	Z28925 HSC1H081.S
34	41.8	2.4	839	82	CNS004NB	AL054280 Drosoph11
35	41.6	2.4	733	82	CNS011AC	AL100014 Drosoph11
36	41.4	2.4	608	44	AW1258757	AW1258757 LP02029.5
37	40.8	2.3	478	24	H73136	H73136 yu03f07.r1
38	40.2	2.3	1101	83	CNS017SY	AL1084460 Drosoph11
39	40	2.3	659	52	AW1651838	AW1651838 wb55h11.x
40	40	2.3	710	52	AW1638881	AW1638881 t108d08.x
41	39.8	2.3	479	81	AW438208	AW438208 707067D05
42	39.8	2.3	486	23	H20378	H20378 yu61d07.r1
43	39.8	2.3	641	81	AW433361	AW433361 707067D05
44	39.2	2.2	409	23	D52750	D52750 HUM086G12B
45	39.2	2.2	964	82	CNS003WG	AL065254 Drosoph11

ALIGNMENTS

RESULT 1
 LOCUS AA459837 503 bp mRNA EST 09-JUN-1997
 DEFINITION zx50g12.r1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:195718
 5' similar to TR:G834328 G834328 SEMAPHORIN C ; mRNA sequence.
 ACCESSION AA459837.1 GI:2184744
 VERSION AA459837.1
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (bases 1 to 503)
 Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S.,
 Kuababa, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J.,
 Moore, B., Scheinberg, K., Steptoe, M., Tan, F., Theising, B.,
 White, Y., Wylie, T., Waterston, R. and Wilson, R.
 T11 JE WashU-Merck EST Project 1997
 JOL RNAL Unpublished (1997)
 COMMENT On Sep 12, 1996 this sequence version replaced gi:1392936.

Contact: Wilson RK
 Washington University School of Medicine
 444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@wustl.wustl.edu
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.lnl.gov) for further information.
 Seq primer: -28m13 rev2 ET from Amersham
 High quality sequence stop: 459.
 Location/Qualifiers
 1..503
 /organism="Homo sapiens"
 /db_xref="GDB:6039163"
 /db_xref="taxon:9606"
 /clone="IMAGE:795718"
 /clone_1db="Soares_testis_NHT"
 /sex="male"
 /lab_host="DH10B"
 /note="Vector: pT73D-Pac (Pharmacia) with a modified
 polylinker. Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
 was prepared from mRNA obtained from Clontech
 Laboratories, Inc., and primed with a Not I - oligo(dT)
 primer [5']
 TGTTCACATCTGAGTGGAGGCGGCCCAATTTTCTTTTCTTTT 3').
 Double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Not I and cloned into the Not I
 and Eco RI sites of the modified pT73D vector. Library
 went through one round of normalization to Colb, and was
 constructed by Bento Soares and M. Fatima Bonaldo. "

BASE COUNT 105 a 136 c 149 g 113 t
 ORIGIN
 Query Match 27.1%; Score 476.4; DB 34; Length 503;
 Best Local Similarity 99.4%; Pred. No. 1.3e-117;
 Matches 499; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

748 gctcttcacagatcgttgcacacaggttgacacagctctcagggaaagatgatgtg 807
 1 GCCTTCTCAGAGTGTGTCGCCACA-GGTGACACACCTCTCAGGAAAGATGATGTG 59
 808 cctcaccgtgggaca-gaggatgagacacctccacacagcagtgagtcagtcagct 866
 60 CTCTGCTGGGACATGAGATGACACTCCACGACATGCGATCGGATCGACT 119
 867 cagcgtctctgagaatctgagccttattccacagcagcagtcagtcagtcagtc 926
 120 CAGCGTCTTGAAGATCTGCGCTTATCCAGACGACACGATGAGACATGAAATT 179
 927 gtaccacagcgtgctcctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 986
 180 GTACACAGCTGCTCCTGTTGCTCCGCTACGATGAGTGACACAGTAAATACACAA 239
 987 ctgtgacgtctcagacgtctcagagtgcatctcagtcagtcagtcagtcagtcagtc 1046
 240 CTGTGCGGCTCTCCAGACTGCTGAGTGCATCTGCGCCACGACCAAGCTGTGCTG 299
 1047 gagcttcagcgtgagtgagtggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1106
 300 GAGCTTCCGCGCTGATGAGTGTGTCCTGAGTCCGCGGACGACGAGGTTGTCAAGA 359
 1107 catagagtcagcagatgtctctcttctgtgtcttaagaagcttgagagtcagtcagtc 1166
 360 CATAGAGTCAGCAGTGTCTCTTGTGTCTTGAAGAGCTTGAGACGCTCCAGTGTGT 419
 1167 gtttaagctccgt 1226
 420 GTTTAAATTCCTCCGCTACAGCTGCGGACATGTGTCTTGCATGTCTTCCAAAGTCAC 479
 1227 atggcaccctctgt 1248
 480 ATGGGACATCTGT 501

[illegible]

LOCUS	517 bp	mRNA	EST	18-MAY-1995
DEFINITION	R54387.1	Soares infant brain INIB Homo sapiens CDNA clone IMAGE:39491.5	similar to SP:A49069 A49069 COLLABSIN -	mRNA
ACCESSION	R54387			
VERSION	R54387.1	GI:816289		
KEYWORDS	EST.			
SOURCE	human.			
ORGANISM	Homo sapiens			
REFERENCE	Eukariota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
AUTHORS	1 (bases 1 to 517) Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevaaskis, E., Waterston, R., Williamson, A., Wohlmann, P. and Wilson, R.			
TITLE	The Washu-Merck EST Project			
JOURNAL	Unpublished (1995)			
COMMENT	Contact: Wilson RK Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: est@watson.wustl.edu Insert Size: 2134 High quality sequence stops: 99 Clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.lnl.gov) for further information. Insert Length: 2134 Std Error: 0.00 Seq primer: M13RPI High quality sequence stop: 99. location/Qualifiers 1. 517 /organism="Homo sapiens" /db_xref="GDB:412032" /db_xref="taxon:9606" /clone="IMAGE:39491" /clone_id="Soares infant brain INIB" /sex="female" /dev_stage="73 days post natal" /lab_host="DH10B (ampicillin resistant)" /note="Organ: whole brain; Vector: lrfm1d BA; Site:1: Not I; Site:2: Hind III; 1st strand cDNA was primed with a Not I - Oligo(dT) primer [5 AACGAGAGATTGCGCGCCGACAGAAATTTTTTTTTTTT 3']; double-stranded cDNA was ligated to Hind III adaptors (Pharmacia), digested with Not I and directionally cloned into the Not I and Hind III sites of the lrfm1d BA vector. Library went through one round of normalization. Library constructed by Bento Soares and M.Fatima Bonaldo."			
BASE COUNT	123 a 137 c 128 g 118 t			
ORIGIN				
Query Match	19.88	Score 349.2	DB: 22	Length 517;
Best Local Similarity	95.18	Pred. No. 2.1e-83		
Matches 411: Conservative	0	Mismatches 14	Indels 7	Gaps 5;
488	gaccacaagaacatcgcagcagctgcgaatggtcccttcagaagaactaaacatgactgca	547		
Db	1 GACCACAGACATTCGGACAGTGGATGATGATCCCTTCAGAGAACTAAACATGACTGCA	60		
548	acagagagactcctcgctgctgtgacaaatgatgtgtcccaagccagacactgagatgacatca	607		
Db	61 ACAGGAGACTCCTGCTGCTGACATGATGTGTGCCACCCAGACCTTGAGAGATGCATCA	120		
608	ccaacaataaagactcgcgcgacacttgatgcatctctctccctgcctcagcgcgtaccca	667		
Db	121 CCACCAATATGAAGCTCGCGACATTGGCTATCTCTCTCCCTGCTGACCGCGTACTCA	180		
668	ccttcatacgcggaccacccactatgacacgagccagtgatcttcacagtatgagccacccc	727		

DB	181	CCITCATCGGGGANGACCACCATCAGTACAGGACAGGAGCAGTCTTTTCAGCTGATGATGCCACCC	240
OY	728	ttcgtgcaactacagatacaagcattctcgaagtcgtgtggcccaagggtgacagcctct	787
DB	241	TGNTGCTACTACATACATACAGNCTATCTTCACAGAGTGTGGCCACAGAGGTGACAGCCTCT	300
OY	788	cagggaagaagatgatattgctctaccct-gggggaagaagat-ggacaccaccacagagc	845
DB	301	CAGGGAAGAGATATATGCTGCTACCTGGGGGAGAGAGATGGGCAACTTTCACCCGAGC	360
OY	846	agtgaggatcaggagctcaagctcagcgtt-cttgaagatct-ggccttatcccaag	900
DB	361	AGTGGGATTCGACACTCAGTTCACAGCTTCTTTCGAATCTGGCCTATTTCACAG	420
OY	901	ccacagccagct 912	
DB	421	TCACAGCAGT 432	
RESULT 4			
245329	337 bp	mRNA	EST
LOCUS	HS22MC091	normalized infant brain cDNA	Homo sapiens cDNA clone
DEFINITION	C-2mg09, mRNA sequence.		
ACCESSION	245329		
VERSION	245329.1	GI:574541	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;		
	Eutheria; Primates; Catarrhini; Homnidae; Homo.		
REFERENCE	1 (bases 1 to 337)		
AUTHORS	Aufrey,C., Behar,G., Bois,F., Bouchier,C., da Silva,C., Devignes,M.D., Duprat,S., Houlgatte,R., Jumeau,M.N., Lamy,B., Lorenzo,F., Mitchell,H., Marlage-Samson,R., Pietu,G., Pouliot,Y., Sebastiant-Rebatachis,C. and Tessier,A. IMAGE: molecular integration of the analysis of the human genome and its expression C. R. Acad. Sci. III, Sci. Vie 318 (2), 263-272 (1995)		
TIJLE	95277534		
JOURNAL	Contact: Genethon		
MEJINE	Genethon Centre de recherche sur le Genome Humain		
COMMENT	1,rue de l'Internationale, Bp60 91002 Evry Cedex, FRANCE		
	Tel: 33169472800		
	Fax: 33160778698		
	Email: genexpress@genethon.fr		
	Single read.		
	Genexpress_library_id: C; Genexpress_sequence_id: ylc-2mg09		
	Seq primer: (-21)M3-universal.		
FEATURES	Location/Qualifiers		
SOURCE	1..337		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="C-2mg09"		
	/clone_lib="normalized infant brain cDNA"		
	/sex="Female"		
	/tissue_type="total brain"		
	/dev_stage="3 months old"		
	/note="Organ: brain; Vector: lafmid BA; Site1: HindIII;		
	Site2: NotI; sex:Female; dev_stage=3 months old;		
	isolate=muscular atrophy patient; tissue_type=total		
	brain; total mRNA was Oligo-(dT) primed and directional		
	cloned 5' -> 3' into the HindIII -> NotI sites of the		
	laifmid B vector. Clone library from B.Scares, Psychia-		
	Dep. Columbia University, USA. Normalization_method:		
	Bento Scares, P.N.A.S in press"		
BASE COUNT	81 a	106 c	82 g
ORIGIN	67 t	1	others

Matches	335;	Conservative	0;	Mismatches	2;	Indels	0;	Gaps	0;
QY	487	cgaccacaagacatctcgagacagtgctgaatggtcccttcagagaactaaacatgactgc	546						
Db	1	CGACCACAGACATCTCGGACAGTGCCTGATGTCCTCTCAGAGAACTAAACATGACTGC	60						
QY	547	aacagaagacatgcctgtgttgacataatgtgtcccccagaccctggagagtc	606						
Db	61	AACAGAGGCTGCGCTGTGTGTGTGACATATGTCGCCACAGCCAGACTGTGAGATGCATC	120						
QY	607	accacaacatgaaagctcccgacacttggctcatctctctcccttcgaccgcgtactc	666						
Db	121	ACCACACATGATGAACCTCGGACCTTGTGGCTCATGTCTCTCCTCTCCGACCGCTACTC	180						
QY	667	acctcatcccgagaccaccacactcatgtgacagcgccagtgttccagctgatagtccccc	726						
Db	181	ACCTTCATCTCCGGGACCCACCTCATGTGACAGGCGCAAGTGTTCAGCTGATGGCCACCCC	240						
QY	727	ctgctggtcactacagatatacagccatcttcagagtcgtggtggccaaaggtgaaccagctc	786						
Db	241	CTGCTGTCCTACTACAGTTACAGCTCATCTCAGAGCTGTGGCCACAGGGGTACACGCTC	300						
QY	787	tcaaggaaagatgatagtctctctacctcgaggagac	823						
Db	301	TCAGGAAAGATGATGATGTCTCTACTCTGGGACAG	337						
RESULT	5								
H24181	437 bp	mRNA	EST	06-JUL-1995					
LOCUS	YM55f02.r1	Soares infant brain INIB	Homo sapiens cDNA clone						
DEFINITION	IMAGE:52280 5' similar to SP:A49069 A45069 COLLASIN - ;, mRNA								
ACCESSION	H24181								
VERSION	H24181.1	GI:892876							
KEYWORDS	EST.								
SOURCE	human.								
ORGANISM	Homo sapiens								
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;								
	Eutheria; Primates; Catarrhini; Homnidae; Homo.								
REFERENCE	1 (bases 1 to 437)								
AUTHORS	Hallier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,								
	Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,								
	Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,								
	Trevaskis,E., Waterston,R., Williamson,A., Wohlmann,P. and								
	Wilson,R.								
TITLE	The Mashu-Merck EST Project								
JOURNAL	Unpublished (1995)								
COMMENT	Contact: Wilson RK								
	Washington University School of Medicine								
	444 Forest Park Parkway, Box 8501, St. Louis, MO 63108								
	Tel: 314 286 1800								
	Fax: 314 286 1810								
	Email: est@watson.wustl.edu								
	Insert Size: 1989								
	High quality sequence stops: 310								
	This clone is available royalty-free through LNL; contact the								
	IMAGE Consortium (info@image.llnl.gov) for further information.								
	Insert length: 1989 Std Error: 0.00								
	Seq primer: M13RP1								
	High quality sequence stop: 310.								
FEATURES	Location/Qualifiers								
source	1. 437								
	/organism="Homo sapiens"								
	/db_xref="GDB:425216"								
	/db_xref="taxon:9606"								
	/clone="IMAGE:52280"								
	/clone_id="Soares infant brain INIB"								
	/sex="female"								
	/dev_stage="73 days post natal"								

I - oligo(dT) primer [5';
 ACTGAGAAATTCGGCGCCGACAGATTTTCTTTTCTTTT 3'];
 double-stranded cDNA was ligated to Hind III adaptors
 (Pharmacia), digested with Not I and directionally cloned
 into the Not I and Hind III sites of the Lactid BA vector.
 Library went through one round of normalization. Library
 constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 105 a 122 c 112 g 95 t 3 others

Query Match 18.6%; Score 328.4; DB 23; Length 437;
 Best Local Similarity 97.5%; Pred. No. 7.4e-78;
 Matches 344; Conservative 0; Mismatches 7; Indels 2; Gaps 1;

OY 488 gaccacaagacatcgcgcagctgcaatgctccctcagagaactaaacatgactga 547
 |||||||
 Db 1 GACCAAGAGACTTGGACAGTCGATGCTCCTTCAGAACTAAACATGACTGCA 60
 OY 548 acagagagactgcctgctgagacaatgctgccccagaccagactggaagatgcatca 607
 |||||||
 Db 61 ACAGAGGAGACTGCTGCTGAGCAATGATGTCGCCAGGCCAGACCTGAGAGTGCATCA 120
 OY 608 ccaacaacatgaagctcgcgcaccttgctcactctcctccctgactgaaccgactactca 667
 |||||||
 Db 121 CCACACATATGAGCTCGGCGACCTTGGCTCATCTCTCCTGCTGACCGGCTACTCA 180
 OY 668 cctcatccgcgcagacacacacatcagacagcagctgttccagctgagccaccccc 727
 |||||||
 Db 181 CCTTATCCGGGACACCCACATGAGAGCCAGTGTTCACCTGATGAGCCACCC 240
 OY 728 tgcctgactacacagatacagcctatcagagctgctgagccagaggtgacagcctct 787
 |||||||
 Db 241 TGCTGCTACTACATACAGCTTACGCTTCTCAGAGTGTGCCCCACAGAGGAGACACCTTT 300
 OY 788 cagggaaagatgatgctgctct--acctggagacagagatgacacctctc 838
 |||||||
 Db 301 TAGGAAAGAGATGATGCTGCTTACTCTGGGGACAGAGATGAGACACTTTC 353

RESULT 6
 H10623 347 bp mRNA EST 23-JUN-1995
 LOCUS ym8907.r1 Soares Infant Brain INIB Homo sapiens cDNA clone
 DEFINITION IMAGE:47298 5' similar to SP:A49069 A49069 COLLABSIN - ; mRNA
 sequence.

ACCESSION H10623
 VERSIONS H10623.1 GI:875445
 KEYWORDS EST
 SOURCE human.

ORANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Homiidae; Homo.
 1 (bases 1 to 347)
 Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,
 Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,
 Paterson, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F.,
 Tevaskis, E., Waterston, R., Williamson, A., Wohlmann, P. and
 Wilson, R.
 The Mashu-Merck EST Project
 Unpublished (1995)
 On May 10, 1995 this sequence version replaced gi:805665.
 Contact: Wilson RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@wustl.edu
 Insert Size: 1985
 High quality sequence stops: 198
 Source: IMAGE Consortium, LNL
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.lnl.gov) for further information.

TI LE
 JO RNAL
 COMM NT

Insert Length: 1985 Std Error: 0.00
 Seq Primer: M13RPI
 High quality sequence stop: 198.
 Location/Qualifiers
 1. 347
 /organism="Homo sapiens"
 /db_xref="GDB:419835"
 /db_xref="taxon:9606"
 /clone="IMAGE:47298"
 /clone_id="Soares Infant Brain INIB"
 /sex="female"
 /dev_stage="73 days post natal"
 /lab_host="DH10B (ampicillin resistant)"
 /note="Organ: whole brain; Vector: Lactid BA; Site: 1: Not
 I; Site 2: Hind III; 1st strand cDNA was primed with a Not
 I - oligo(dT) primer [5';
 ACTGAGAAATTCGGCGCCGACAGATTTTCTTTTCTTTT 3'];
 double-stranded cDNA was ligated to Hind III adaptors
 (Pharmacia), digested with Not I and directionally cloned
 into the Not I and Hind III sites of the Lactid BA vector.
 Library went through one round of normalization. Library
 constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 82 a 105 c 87 g 6 t 6 others

Query Match 18.5%; Score 326.4; DB 22; Length 347;
 Best Local Similarity 98.5%; Pred. No. 2.3e-77;
 Matches 327; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 488 gaccacaagacatcgcgcagctgcaatgctccctcagagaactaaacatgactga 547
 |||||||
 Db 1 GACCAAGAGACTTGGACAGTCGATGCTCCTTCAGAACTAAACATGACTGCA 60
 OY 548 acagagagactgcctgctgagacaatgctgccccagaccagactggaagatgcatca 607
 |||||||
 Db 61 ACAGAGGAGACTGCTGCTGAGCAATGATGTCGCCAGGCCAGACCTGAGAGTGCATCA 120
 OY 608 ccaacaacatgaagctcgcgcaccttgctcactctcctccctgactgaaccgactactca 667
 |||||||
 Db 121 CCACACATATGAGCTCGGCGACCTTGGCTCATCTCTCCTGCTGACCGGCTACTCA 180
 OY 668 cctcatccgcgcagacacacacatcagacagcagctgttccagctgagccaccccc 727
 |||||||
 Db 181 CCTTATCCGGGACACCCACATGAGAGCCAGTGTTCACCTGATGAGCCACCC 240
 OY 728 tgcctgactacacagatacagcctatcagagctgctgagccagaggtgacagcctct 787
 |||||||
 Db 241 TGCTGCTACTACATACAGCTTACGCTTCTCAGAGTGTGCCCCACAGAGGAGACACCTTT 300
 OY 788 cagggaaagatgatgctgctctcactcctgag 819
 |||||||
 Db 301 CAGGNAAGAGATGATGCTGCTTACTCTGGGG 332

RESULT 7
 AM436751 376 bp mRNA EST 14-FEB-2000
 LOCUS 77209 MARC 2P1G Sus scrofa cDNA 5', mRNA sequence.
 DEFINITION AM436751
 ACCESSION AM436751
 VERSION AM436751.1 GI:6972057
 KEYWORDS EST
 SOURCE Sus scrofa
 pig.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Cetartiodactyla; Suidae; Suidae; Sus.
 1 (bases 1 to 376)
 Fahnenkrug, S.C., Freking, B.A., Rohrer, G.A., Smith, T.P.L., Casas, E.,
 Stone, R.T., Heaton, M.P., Grosse, W.M., Bennett, G.A. and Keele, J.W.
 Design and use of two pooled tissue normalized cDNA libraries for
 EST discovery in swine
 Unpublished (2000)
 On Jul 9, 1999 this sequence version replaced gi:5435515.

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

RESU T 9
AM436702 208 bp mRNA EST 14-FEB-2000
LOCUS: 77141 MARC 2Pig Sus scrofa cDNA 5', mRNA sequence.
DEFINITION: AM436702
ACCESSION: AM436702
VERSION: EST
KEYWORDS: GI:6972008
SOURCE: Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
ORIGIN: Faintkrug, S.C., Freking, B.A., Rohrer, G.A., Smith, T.P.L., Casas, E., Stone, R.T., Heath, M.P., Grosse, W.M., Bennett, G.A. and Keeler, J.W. Design and use of two pooled tissue normalized cDNA libraries for EST discovery in swine
Unpublished (2000)
On Jul 9, 1999 this sequence version replaced gi:5435372.
COMMENT: Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@mail.marc.usda.gov
Single pass sequencing. Bases called and trimmed with phred v0.980904.e. Vector identified by cross_match with the -mnscore 20 and -mismatch 12 options.
PCR Primers
FORWARD: AGGAACAGCTATGACAT
BACKWARD: GTTCCACACGACG
Plate: 32 row: E column: 18
Seq primer: ATTAGTGACACTATAG.
Location/Qualifiers
1..208
/organism="Sus scrofa"
/db_xref="taxon:9823"
/clone_1lb="MARC 2Pig"
/tissue_type="pooled"
/lab_host="DH10B"
/note="Vector: PCMV SPORT6; site_1: xbaI; site_2: XhoI; library made from pooled tissue from testis, ovary, endometrium, hypothalamus, pituitary, and placenta."
BASE COUNT 36 a 64 c 73 g 35 t
ORIGIN

Query Match 9.5%; Score 167; DB 81; Length 208;
Best Local Similarity 87.9%; Pred. No. 1,2e-34;
Matches 182; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

QY 1393 agcgaagagatgctcgaagccgagccacacagatggggcgagagctggtctcttc 1452
|||||
DB 2 AGCCACCGGGGGCCCCAGAGCCGGGCTCACAGTGGGGCCGAGTGGCAGCGCTTCTTC 61
QY 1453 ttgggagattctgcagcattccctgactctcatcttattgttcggtcgaagcgagcg 1512
|||||
DB 62 CTGGGGGTTCTTGACATCCCTGACTCTTCTCTATTGGTGGGCTGACGCGGCAg 121
QY 1513 ccagaaaggaactcttgctagaagacaaggtggcgctgagctggggctccactctc 1572
|||||
DB 122 CGACAGAGGAGGAGTGTGGCTAGAGACAAAGTGGGCTTGAGCTGGGCTCCAGCATCC 181
QY 1573 gggaaccacaagctacagccaagacccct 1599
|||||
DB 182 GGGACCACAGCTACAGCCAGGACCT 208
RESULT 10
LOCUS: 242773 212 bp mRNA EST 11-NOV-1994
DEFINITION: HSC07E021 normalized infant brain cDNA Homo sapiens cDNA clone
SOURCE: c-07e02, mRNA sequence.

ACCESSION 242773
VERSION 242773.1 GI:570345
KEYWORDS: EST.
SOURCE: human.
ORGANISM: Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE: 1 (bases 1 to 212)
Auffray, C., Behar, G., Bols, F., Bouchler, C., da Silva, C., Devignes, M.D., Duprat, S., Houligatte, R., Juneau, M.N., Lamy, B., Lorenzo, F., Mitchell, H., Mariage-Samson, R., Pletu, G., Pouillot, Y., Sebastiani-Kabakchis, C. and Tessier, A.
IMAGE: molecular integration of the analysis of the human genome and its expression
C. R. Acad. Sci. III, Sci. Vie 318 (2), 263-272 (1995)
95277534
JOURNAL MEDLINE
TITLE: Contact: Genethon
Genethon Centre de recherche sur le Genome Humain
1, rue de l'Internationale, BP60 91002 EVRY Cedex, FRANCE
Tel: 33169472800
Fax: 33160778698
Email: genexpress@genethon.fr
Single read.
Genexpress_library_id: C; Genexpress_sequence_id: y1c-07e02
Seq primer: (-21)M13-universal.
Location/Qualifiers
1..212
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_1lb="c-07e02"
/clone_1lb="normalized infant brain cDNA"
/sex="female"
/tissue_type="total brain"
/dev_stage="3 months old"
/note="Organ: brain; Vector: lafmid BA; Site_1: HindIII; Site_2: NotI; sex: female; dev_stage: 3 months old; isolate-muscular atrophy patient; tissue_type-total brain; local mRNA was oligo-(dT) primed and directionally cloned 5' -> 3' into the HindIII -> NotI sites of the lafmid BA vector. Clone library from B.Saeres, Psychiatry Dept. Columbia University, USA. Normalization_method: Bento Soares, P.N.A.S. in press"
BASE COUNT 48 a 56 c 58 g 45 t 5 others
ORIGIN

Query Match 8.0%; Score 141; DB 20; Length 212;
Best Local Similarity 97.6%; Pred. No. 1,2e-27;
Matches 163; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

QY 835 ctccacagagcagtcgagatcgagctcagctcagctcttgaaagatcggcctatc 894
|||||
DB 28 CTCACCGAGCAGCTGCGAGTCCGAGCTCAG-TCAGGCTTCTGAAGATCTGCGCTTNTN 86
QY 895 ccagagcccaagcagctttagaacaatgaatgttaccagactgtctctgtttgtctc 954
|||||
DB 87 CCAGAGCCACACCCAGTTGAGAACATGAATGTACACAGCTGGCTCTGTTGGCTCC 146
QY 955 cgtactgaagtgacacaagtgatacaacacactgtgctgtccca 1001
|||||
DB 147 CGTACGAGGTGACACACAGTGAATA-AACCACTGTGGCGCTCTCA 192
RESULT 11
LOCUS: AM346693 232 bp mRNA EST 01-FEB-2000
DEFINITION: 29273 MARC 1Pig Sus scrofa cDNA 5', mRNA sequence.
ACCESSION: AM346693
VERSION: AM346693.1 GI:6844403
KEYWORDS: EST.
SOURCE: pig.
ORGANISM: Sus scrofa

THIS PAGE BLANK (USPTO)